STIC-Biotech/ChemLib 160423

From: Hamud, Fozia
Sent: Monday, July 25, 2005 5:44 PM
To: STIC-Biotech/ChemLib
Subject: sequence search for 09/598,443

Please search SEQ ID NOs:1 and 2 of 09/598,443 against commercial and interference data bases. thanks

Fozia Hamud
Patent Examiner
Remsen 4D64
Mail Box: Remsen 4C70
(571) 272-0884
Art Unit 1647

****
STAFF USE ONLY

Searcher:_____
Searcher Phone: 2-
Date Searcher Picked up: 7/26/05
Date Completed:_____
Searcher Prep/Rev. Time:__
Online Time:_____

****
Type of Search

NA#:____/ AA#:____/
Interference:___ SPDI:___
S/L:_____ Oligomer:____
Encode/Transl:_____
Structure#:_____ Text:___
Inventor:____ Litigation:___

****
Vendors and cost where applicable
STN:_____
DIALOG:_____
QUESTEL/ORBIT:_____
LEXIS/NEXIS:_____
SEQUENCE SYSTEM: 0.5 H
WWW/Internet:_____
Other(Specify):_____

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: July 29, 2005, 20:14:15 ; Search time 5659 Seconds
(without alignments)
10557.569 Million cell updates/sec

Title: US-09-598-443-1
Perfect score: 1233
Sequence: 1 atgccaggtgtctgtgatag.....tgtccaaggatgatatgtag 1233

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1233	100.0	1233	6	BD103586	BD103586 SIGIRR DN
	2	1233	100.0	1618	9	BC025953	BC025953 Homo sapi
	3	1233	100.0	1627	9	BC003591	BC003591 Homo sapi
	4	1233	100.0	1657	9	AK172830	AK172830 Homo sapi
	5	1233	100.0	1695	9	AK025099	AK025099 Homo sapi
	6	1231.4	99.9	1659	6	AX058610	AX058610 Sequence
	7	1230.4	99.8	1233	9	CR457338	CR457338 Homo sapi
	8	1071.8	86.9	1649	6	AX201368	AX201368 Sequence
	9	1071.8	86.9	1649	9	AY358342	AY358342 Homo sapi
	10	860.2	69.8	1568	10	BC010806	BC010806 Mus muscu
	11	856.6	69.5	1421	10	AF239957	AF239957 Mus muscu
	12	826	67.0	1540	10	AF113795	AF113795 Mus muscu
	13	556.2	45.1	1867	9	AK093427	AK093427 Homo sapi
c	14	398	32.3	655	11	BV177172	BV177172 sqnm93226
	15	391.6	31.8	641	6	AX525866	AX525866 Sequence
	16	329.6	26.7	915	6	CQ717444	CQ717444 Sequence
	17	265.6	21.5	425	6	AR423041	AR423041 Sequence
	18	265.6	21.5	425	6	AX983735	AX983735 Sequence
	19	265.6	21.5	425	6	BD118594	BD118594 EST and e
c	20	250.6	20.3	557	6	AX400401	AX400401 Sequence
	21	245.4	19.9	249	6	AX894237	AX894237 Sequence
	22	245.4	19.9	249	6	BD029770	BD029770 Sequence
c	23	207.2	16.8	165000	9	AC137894	AC137894 Homo sapi
c	24	207.2	16.8	172449	2	AC142165	AC142165 Homo sapi
	25	207.2	16.8	175416	2	AP006286	AP006286 Homo sapi
c	26	207.2	16.8	187160	9	AC138230	AC138230 Homo sapi
c	27	195.6	15.9	175416	2	AP006286	AP006286 Homo sapi
c	28	186	15.1	388	6	AX408731	AX408731 Sequence
	29	173.4	14.1	61320	2	AC135992	AC135992 Homo sapi
c	30	164.2	13.3	61320	2	AC135992	AC135992 Homo sapi
	31	141.6	11.5	202996	2	AC109272	AC109272 Mus muscu
c	32	141.6	11.5	211591	2	AC108908	AC108908 Mus muscu
	33	134.4	10.9	237066	2	AC134080	AC134080 Rattus no
c	34	134.4	10.9	295957	2	AC122619	AC122619 Rattus no
c	35	122	9.9	324	11	G38549	G38549 SHGC-64262
	36	71.4	5.8	248718	10	AC117629	AC117629 Mus muscu
c	37	69.8	5.7	259	6	AX524228	AX524228 Sequence
c	38	69.8	5.7	259	6	AX552966	AX552966 Sequence
c	39	68.2	5.5	157152	10	AC114679	AC114679 Mus muscu
	40	68.2	5.5	174139	2	AC116700	AC116700 Mus muscu
	41	68	5.5	125020	9	AF429315	AF429315 Homo sapi
c	42	67	5.4	124244	10	AL929026	AL929026 Mouse DNA
	43	67	5.4	124795	10	AL840624	AL840624 Mouse DNA
c	44	66.6	5.4	187126	10	AC113315	AC113315 Mus muscu
c	45	66.2	5.4	116887	10	AL929153	AL929153 Mouse DNA

ALIGNMENTS

RESULT 1
BD103586
LOCUS BD103586 1233 bp DNA linear PAT 27-AUG-2002
DEFINITION SIGIRR DNA and polypeptide.
ACCESSION BD103586
VERSION BD103586.1 GI:22649160
KEYWORDS JP 2001526892-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1233)
AUTHORS Sims,J.E.
TITLE SIGIRR DNA and polypeptide
JOURNAL Patent: JP 2001526892-A 1 25-DEC-2001;
IMMUNEX CORP
COMMENT OS Homo sapiens (human)
PN JP 2001526892-A/1
PD 25-DEC-2001
PF 23-DEC-1998 JP 2000525545
PR 23-DEC-1997 US 60/068770
PI JOHN ERNEST SIMS
PC C12N15/09,C07K14/715,C07K16/28,C12N1/15,C12N1/19,C12N1/21, PC C12N5/10,
PC C12P21/02,C12P21/08,C12Q1/37,G01N27/26,G01N27/447,C12N15/00,
PC C12N5/00,
PC G01N27/26
CC SIGIRR DNA and polypeptide
FH Key Location/Qualifiers
FT source 1. .1233
FT /organism='Homo sapiens (human)'.
FEATURES Location/Qualifiers
source 1. .1233
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN

Query Match 100.0%; Score 1233; DB 6; Length 1233;
Best Local Similarity 100.0%; Pred. No. 3.4e-182;
Matches 1233; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATGCCAGGTGTCTGTGATAGGGCCCCTGACTTCCTCTCCCCGTCTGAAGACCAGGTGCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCCAGGTGTCTGTGATAGGGCCCCTGACTTCCTCTCCCCGTCTGAAGACCAGGTGCTG 60

Qy         61 AGGCCTGCCTTGGGCAGCTCAGTGGCTCTGAACTGCACGGCTTGGGTAGTCTCTGGGCCC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGGCCTGCCTTGGGCAGCTCAGTGGCTCTGAACTGCACGGCTTGGGTAGTCTCTGGGCCC 120

Qy        121 CACTGCTCCCTGCCTTCAGTCCAGTGGCTGAAAGACGGGCTTCCATTGGGAATTGGGGGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CACTGCTCCCTGCCTTCAGTCCAGTGGCTGAAAGACGGGCTTCCATTGGGAATTGGGGGC 180

Qy        181 CACTACAGCCTCCACGAGTACTCCTGGGTCAAGGCCAACCTGTCAGAGGTGCTTGTGTCC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CACTACAGCCTCCACGAGTACTCCTGGGTCAAGGCCAACCTGTCAGAGGTGCTTGTGTCC 240

Qy        241 AGTGTCCTGGGGGTCAACGTGACCAGCACTGAAGTCTATGGGGCCTTCACCTGCTCCATC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGTGTCCTGGGGGTCAACGTGACCAGCACTGAAGTCTATGGGGCCTTCACCTGCTCCATC 300

Qy        301 CAGAACATCAGCTTCTCCTCCTTCACTCTTCAGAGAGCTGGCCCTACAAGCCACGTGGCT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CAGAACATCAGCTTCTCCTCCTTCACTCTTCAGAGAGCTGGCCCTACAAGCCACGTGGCT 360

Qy        361 GCGGTGCTGGCCTCCCTCCTGGTCCTGCTGGCCCTGCTGCTGGCCGCCCTGCTCTATGTC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCGGTGCTGGCCTCCCTCCTGGTCCTGCTGGCCCTGCTGCTGGCCGCCCTGCTCTATGTC 420

Qy        421 AAGTGCCGTCTCAACGTGCTGCTCTGGTACCAGGACGCGTATGGGGAGGTGGAGATAAAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AAGTGCCGTCTCAACGTGCTGCTCTGGTACCAGGACGCGTATGGGGAGGTGGAGATAAAC 480

Qy        481 GACGGGAAGCTCTACGACGCCTACGTCTCCTACAGCGACTGCCCCGAGGACCGCAAGTTC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GACGGGAAGCTCTACGACGCCTACGTCTCCTACAGCGACTGCCCCGAGGACCGCAAGTTC 540

Qy        541 GTGAACTTCATCCTAAAGCCGCAGCTGGAGCGGCGTCGGGGCTACAAGCTCTTCCTGGAC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GTGAACTTCATCCTAAAGCCGCAGCTGGAGCGGCGTCGGGGCTACAAGCTCTTCCTGGAC 600

Qy        601 GACCGCGACCTCCTGCCGCGCGCTGAGCCCTCCGCCGACCTCTTGGTGAACCTGAGCCGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GACCGCGACCTCCTGCCGCGCGCTGAGCCCTCCGCCGACCTCTTGGTGAACCTGAGCCGC 660

Qy        661 TGCCGACGCCTCATCGTGGTGCTTTCGGACGCCTTCCTGAGCCGGGCCTGGTGCAGCCAC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGCCGACGCCTCATCGTGGTGCTTTCGGACGCCTTCCTGAGCCGGGCCTGGTGCAGCCAC 720

Qy        721 AGCTTCCGGGAGGGCCTGTGCCGGCTGCTGGAGCTCACCCGCAGACCCATCTTCATCACC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AGCTTCCGGGAGGGCCTGTGCCGGCTGCTGGAGCTCACCCGCAGACCCATCTTCATCACC 780

Qy        781 TTCGAGGGCCAGAGGCGCGACCCCGCGCACCCGGCGCTCCGCCTGCTGCGCCAGCACCGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TTCGAGGGCCAGAGGCGCGACCCCGCGCACCCGGCGCTCCGCCTGCTGCGCCAGCACCGC 840

Qy        841 CACCTGGTGACCTTGCTGCTCTGGAGGCCCGGCTCCGTGACTCCTTCCTCCGATTTTTGG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CACCTGGTGACCTTGCTGCTCTGGAGGCCCGGCTCCGTGACTCCTTCCTCCGATTTTTGG 900

Qy        901 AAAGAAGTGCAGCTGGCGCTGCCGCGGAAGGTGCGGTACAGGCCGGTGGAAGGAGACCCC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AAAGAAGTGCAGCTGGCGCTGCCGCGGAAGGTGCGGTACAGGCCGGTGGAAGGAGACCCC 960

Qy        961 CAGACGCAGCTGCAGGACGACAAGGACCCCATGCTGATTCTTCGAGGCCGAGTCCCTGAG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CAGACGCAGCTGCAGGACGACAAGGACCCCATGCTGATTCTTCGAGGCCGAGTCCCTGAG 1020

Qy       1021 GGCCGGGCCCTGGACTCAGAGGTGGACCCGGACCCTGAGGGCGACCTGGGTGTCCGGGGG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GGCCGGGCCCTGGACTCAGAGGTGGACCCGGACCCTGAGGGCGACCTGGGTGTCCGGGGG 1080

Qy       1081 CCTGTTTTTGGAGAGCCATCAGCTCCACCGCACACCAGTGGGGTCTCGCTGGGAGAGAGC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CCTGTTTTTGGAGAGCCATCAGCTCCACCGCACACCAGTGGGGTCTCGCTGGGAGAGAGC 1140

Qy       1141 CGGAGCAGCGAAGTGGACGTCTCGGATCTCGGCTCGCGAAACTACAGTGCCCGCACAGAC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CGGAGCAGCGAAGTGGACGTCTCGGATCTCGGCTCGCGAAACTACAGTGCCCGCACAGAC 1200

Qy       1201 TTCTACTGCCTGGTGTCCAAGGATGATATGTAG 1233
              |||||||||||||||||||||||||||||||||
Db       1201 TTCTACTGCCTGGTGTCCAAGGATGATATGTAG 1233
```

RESULT 2
BC025953

```
LOCUS       BC025953               1618 bp    mRNA    linear   PRI 06-OCT-2003
DEFINITION  Homo sapiens single Ig IL-1R-related molecule, mRNA (cDNA clone
            MGC:39141 IMAGE:3939162), complete cds.
ACCESSION   BC025953
VERSION     BC025953.1  GI:19684156
KEYWORDS    MGC.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1618)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
  MEDLINE   22388257
   PUBMED   12477932
REFERENCE   2  (bases 1 to 1618)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-MAR-2002) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Genome Sequence Centre,
            BC Cancer Agency, Vancouver, BC, Canada
            info@bcgsc.bc.ca
            Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
            Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
            Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
            Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
            Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline
            Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott,
            Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
            George Yang, Scott Zuyderduyn, Marco Marra.
```

```
            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAL Plate: 42 Row: b Column: 7
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 11141876.
FEATURES             Location/Qualifiers
     source          1. .1618
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="MGC:39141 IMAGE:3939162"
                     /tissue_type="Lung, small cell carcinoma"
                     /clone_lib="NIH_MGC_7"
                     /lab_host="DH10B-R"
                     /note="Vector: pOTB7"
     gene            1. .1618
                     /gene="SIGIRR"
                     /db_xref="LocusID:59307"
                     /db_xref="MIM:605478"
     CDS             186. .1418
                     /gene="SIGIRR"
                     /codon_start=1
                     /product="SIGIRR protein"
                     /protein_id="AAH25953.1"
                     /db_xref="GI:19684157"
                     /db_xref="LocusID:59307"
                     /db_xref="MIM:605478"
                     /translation="MPGVCDRAPDFLSPSEDQVLRPALGSSVALNCTAWVVSGPHCSL
                     PSVQWLKDGLPLGIGGHYSLHEYSWVKANLSEVLVSSVLGVNVTSTEVYGAFTCSIQN
                     ISFSSFTLQRAGPTSHVAAVLASLLVLLALLLAALLYVKCRLNVLLWYQDAYGEVEIN
                     DGKLYDAYVSYSDCPEDRKFVNFILKPQLERRRGYKLFLDDRDLLPRAEPSADLLVNL
                     SRCRRLIVVLSDAFLSRAWCSHSFREGLCRLLELTRRPIFITFEGQRRDPAHPALRLL
                     RQHRHLVTLLLWRPGSVTPSSDFWKEVQLALPRKVRYRPVEGDPQTQLQDDKDPMLIL
                     RGRVPEGRALDSEVDPDPEGDLGVRGPVFGEPSAPPHTSGVSLGESRSSEVDVSDLGS
                     RNYSARTDFYCLVSKDDM"
     misc_feature    684. .1109
                     /gene="SIGIRR"
                     /note="TIR; Region: TIR domain. The TIR domain is an
                     intracellular signaling domain found in MyD88, interleukin
                     1 receptor and the Toll receptor. Called TIR (by SMART?)
                     for Toll - Interleukin - Resistance"
                     /db_xref="CDD:pfam01582"
ORIGIN

  Query Match             100.0%;  Score 1233;  DB 9;  Length 1618;
  Best Local Similarity   100.0%;  Pred. No. 3.2e-182;
  Matches 1233;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGCCAGGTGTCTGTGATAGGGCCCCTGACTTCCTCTCCCCGTCTGAAGACCAGGTGCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 ATGCCAGGTGTCTGTGATAGGGCCCCTGACTTCCTCTCCCCGTCTGAAGACCAGGTGCTG 245

Qy         61 AGGCCTGCCTTGGGCAGCTCAGTGGCTCTGAACTGCACGGCTTGGGTAGTCTCTGGGCCC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 AGGCCTGCCTTGGGCAGCTCAGTGGCTCTGAACTGCACGGCTTGGGTAGTCTCTGGGCCC 305

Qy        121 CACTGCTCCCTGCCTTCAGTCCAGTGGCTGAAAGACGGGCTTCCATTGGGAATTGGGGGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        306 CACTGCTCCCTGCCTTCAGTCCAGTGGCTGAAAGACGGGCTTCCATTGGGAATTGGGGGC 365

Qy        181 CACTACAGCCTCCACGAGTACTCCTGGGTCAAGGCCAACCTGTCAGAGGTGCTTGTGTCC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        366 CACTACAGCCTCCACGAGTACTCCTGGGTCAAGGCCAACCTGTCAGAGGTGCTTGTGTCC 425

Qy        241 AGTGTCCTGGGGGTCAACGTGACCAGCACTGAAGTCTATGGGGCCTTCACCTGCTCCATC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        426 AGTGTCCTGGGGGTCAACGTGACCAGCACTGAAGTCTATGGGGCCTTCACCTGCTCCATC 485

Qy        301 CAGAACATCAGCTTCTCCTCCTTCACTCTTCAGAGAGCTGGCCCTACAAGCCACGTGGCT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        486 CAGAACATCAGCTTCTCCTCCTTCACTCTTCAGAGAGCTGGCCCTACAAGCCACGTGGCT 545

Qy        361 GCGGTGCTGGCCTCCCTCCTGGTCCTGCTGGCCCTGCTGCTGGCCGCCCTGCTCTATGTC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        546 GCGGTGCTGGCCTCCCTCCTGGTCCTGCTGGCCCTGCTGCTGGCCGCCCTGCTCTATGTC 605

Qy        421 AAGTGCCGTCTCAACGTGCTGCTCTGGTACCAGGACGCGTATGGGGAGGTGGAGATAAAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        606 AAGTGCCGTCTCAACGTGCTGCTCTGGTACCAGGACGCGTATGGGGAGGTGGAGATAAAC 665

Qy        481 GACGGGAAGCTCTACGACGCCTACGTCTCCTACAGCGACTGCCCCGAGGACCGCAAGTTC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        666 GACGGGAAGCTCTACGACGCCTACGTCTCCTACAGCGACTGCCCCGAGGACCGCAAGTTC 725

Qy        541 GTGAACTTCATCCTAAAGCCGCAGCTGGAGCGGCGTCGGGGCTACAAGCTCTTCCTGGAC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        726 GTGAACTTCATCCTAAAGCCGCAGCTGGAGCGGCGTCGGGGCTACAAGCTCTTCCTGGAC 785

Qy        601 GACCGCGACCTCCTGCCGCGCGCTGAGCCCTCCGCCGACCTCTTGGTGAACCTGAGCCGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        786 GACCGCGACCTCCTGCCGCGCGCTGAGCCCTCCGCCGACCTCTTGGTGAACCTGAGCCGC 845

Qy        661 TGCCGACGCCTCATCGTGGTGCTTTCGGACGCCTTCCTGAGCCGGGCCTGGTGCAGCCAC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        846 TGCCGACGCCTCATCGTGGTGCTTTCGGACGCCTTCCTGAGCCGGGCCTGGTGCAGCCAC 905

Qy        721 AGCTTCCGGGAGGGCCTGTGCCGGCTGCTGGAGCTCACCCGCAGACCCATCTTCATCACC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        906 AGCTTCCGGGAGGGCCTGTGCCGGCTGCTGGAGCTCACCCGCAGACCCATCTTCATCACC 965

Qy        781 TTCGAGGGCCAGAGGCGCGACCCCGCGCACCCGGCGCTCCGCCTGCTGCGCCAGCACCGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        966 TTCGAGGGCCAGAGGCGCGACCCCGCGCACCCGGCGCTCCGCCTGCTGCGCCAGCACCGC 1025

Qy        841 CACCTGGTGACCTTGCTGCTCTGGAGGCCCGGCTCCGTGACTCCTTCCTCCGATTTTTGG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1026 CACCTGGTGACCTTGCTGCTCTGGAGGCCCGGCTCCGTGACTCCTTCCTCCGATTTTTGG 1085

Qy        901 AAAGAAGTGCAGCTGGCGCTGCCGCGGAAGGTGCGGTACAGGCCGGTGGAAGGAGACCCC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1086 AAAGAAGTGCAGCTGGCGCTGCCGCGGAAGGTGCGGTACAGGCCGGTGGAAGGAGACCCC 1145

Qy        961 CAGACGCAGCTGCAGGACGACAAGGACCCCATGCTGATTCTTCGAGGCCGAGTCCCTGAG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1146 CAGACGCAGCTGCAGGACGACAAGGACCCCATGCTGATTCTTCGAGGCCGAGTCCCTGAG 1205

Qy       1021 GGCCGGGCCCTGGACTCAGAGGTGGACCCGGACCCTGAGGGCGACCTGGGTGTCCGGGGG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1206 GGCCGGGCCCTGGACTCAGAGGTGGACCCGGACCCTGAGGGCGACCTGGGTGTCCGGGGG 1265

Qy       1081 CCTGTTTTTGGAGAGCCATCAGCTCCACCGCACACCAGTGGGGTCTCGCTGGGAGAGAGC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1266 CCTGTTTTTGGAGAGCCATCAGCTCCACCGCACACCAGTGGGGTCTCGCTGGGAGAGAGC 1325

Qy       1141 CGGAGCAGCGAAGTGGACGTCTCGGATCTCGGCTCGCGAAACTACAGTGCCCGCACAGAC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1326 CGGAGCAGCGAAGTGGACGTCTCGGATCTCGGCTCGCGAAACTACAGTGCCCGCACAGAC 1385

Qy       1201 TTCTACTGCCTGGTGTCCAAGGATGATATGTAG 1233
              ||||||||||||||||||||||||||||||||
Db       1386 TTCTACTGCCTGGTGTCCAAGGATGATATGTAG 1418


RESULT 3
BC003591
LOCUS       BC003591                1627 bp    mRNA    linear   PRI 30-JUN-2004
DEFINITION  Homo sapiens single Ig IL-1R-related molecule, mRNA (cDNA clone
            MGC:4338 IMAGE:2821373), complete cds.
ACCESSION   BC003591
VERSION     BC003591.1  GI:13097794
KEYWORDS    MGC.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1627)
```

AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.

TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

PUBMED 12477932

REFERENCE 2 (bases 1 to 1627)

AUTHORS Strausberg,R.

TITLE Direct Submission

JOURNAL Submitted (20-FEB-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov

COMMENT Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTP
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield, Andy Chan, Steve S. Chand, William Chow, Alison Cloutier, Ruth Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao, Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin, Teika Olson, Diana Palmquist, Anca Petrescu, Anna Liisa Prahbu, Parvaneh Saeedi, JR Santos, Angelique Schnerch, Ursula Skalska, Duane Smailus, Jeff Stott, Miranda Tsai, George Yang, Jacquie Schein, Asim Siddiqui, Rob Holt, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 9 Row: n Column: 5
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 11141876.

FEATURES Location/Qualifiers

source 1. .1627
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:4338 IMAGE:2821373"
/tissue_type="Lung, small cell carcinoma"
/clone_lib="NIH_MGC_7"
/lab_host="DH10B-R"
/note="Vector: pOTB7"

gene 1. .1627
/gene="SIGIRR"
/db_xref="LocusID:59307"
/db_xref="MIM:605478"

CDS 157. .1389
/gene="SIGIRR"
/codon_start=1
/product="single Ig IL-1R-related molecule"
/protein_id="AAH03591.1"
/db_xref="GI:13097795"
/db_xref="LocusID:59307"
/db_xref="MIM:605478"
/translation="MPGVCDRAPDFLSPSEDQVLRPALGSSVALNCTAWVVSGPHCSL
PSVQWLKDGLPLGIGGHYSLHEYSWVKANLSEVLVSSVLGVNVTSTEVYGAFTCSIQN
ISFSSFTLQRAGPTSHVAAVLASLLVLLALLLAALLYVKCRLNVLLWYQDAYGEVEIN
DGKLYDAYVSYSDCPEDRKFVNFILKPQLERRRGYKLFLDDRDLLPRAEPSADLLVNL
SRCRRLIVVLSDAFLSRAWCSHSFREGLCRLLELTRRPIFITFEGQRRDPAHPALRLL
RQHRHLVTLLLWRPGSVTPSSDFWKEVQLALPRKVRYRPVEGDPQTQLQDDKDPMLIL
RGRVPEGRALDSEVDPDPEGDLGVRGPVFGEPSAPPHTSGVSLGESRSSEVDVSDLGS
RNYSARTDFYCLVSKDDM"

ORIGIN

Query Match 100.0%; Score 1233; DB 9; Length 1627;
Best Local Similarity 100.0%; Pred. No. 3.2e-182;
Matches 1233; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATGCCAGGTGTCTGTGATAGGGCCCCTGACTTCCTCTCCCCGTCTGAAGACCAGGTGCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        157 ATGCCAGGTGTCTGTGATAGGGCCCCTGACTTCCTCTCCCCGTCTGAAGACCAGGTGCTG 216

Qy         61 AGGCCTGCCTTGGGCAGCTCAGTGGCTCTGAACTGCACGGCTTGGGTAGTCTCTGGGCCC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        217 AGGCCTGCCTTGGGCAGCTCAGTGGCTCTGAACTGCACGGCTTGGGTAGTCTCTGGGCCC 276

Qy        121 CACTGCTCCCTGCCTTCAGTCCAGTGGCTGAAAGACGGGCTTCCATTGGGAATTGGGGGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        277 CACTGCTCCCTGCCTTCAGTCCAGTGGCTGAAAGACGGGCTTCCATTGGGAATTGGGGGC 336

Qy        181 CACTACAGCCTCCACGAGTACTCCTGGGTCAAGGCCAACCTGTCAGAGGTGCTTGTGTCC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        337 CACTACAGCCTCCACGAGTACTCCTGGGTCAAGGCCAACCTGTCAGAGGTGCTTGTGTCC 396

Qy        241 AGTGTCCTGGGGGTCAACGTGACCAGCACTGAAGTCTATGGGGCCTTCACCTGCTCCATC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        397 AGTGTCCTGGGGGTCAACGTGACCAGCACTGAAGTCTATGGGGCCTTCACCTGCTCCATC 456

Qy        301 CAGAACATCAGCTTCTCCTCCTTCACTCTTCAGAGAGCTGGCCCTACAAGCCACGTGGCT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        457 CAGAACATCAGCTTCTCCTCCTTCACTCTTCAGAGAGCTGGCCCTACAAGCCACGTGGCT 516

Qy        361 GCGGTGCTGGCCTCCCTCCTGGTCCTGCTGGCCCTGCTGCTGGCCGCCCTGCTCTATGTC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        517 GCGGTGCTGGCCTCCCTCCTGGTCCTGCTGGCCCTGCTGCTGGCCGCCCTGCTCTATGTC 576

Qy        421 AAGTGCCGTCTCAACGTGCTGCTCTGGTACCAGGACGCGTATGGGGAGGTGGAGATAAAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        577 AAGTGCCGTCTCAACGTGCTGCTCTGGTACCAGGACGCGTATGGGGAGGTGGAGATAAAC 636

Qy        481 GACGGGAAGCTCTACGACGCCTACGTCTCCTACAGCGACTGCCCCGAGGACCGCAAGTTC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        637 GACGGGAAGCTCTACGACGCCTACGTCTCCTACAGCGACTGCCCCGAGGACCGCAAGTTC 696

Qy        541 GTGAACTTCATCCTAAAGCCGCAGCTGGAGCGGCGTCGGGGCTACAAGCTCTTCCTGGAC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        697 GTGAACTTCATCCTAAAGCCGCAGCTGGAGCGGCGTCGGGGCTACAAGCTCTTCCTGGAC 756

Qy        601 GACCGCGACCTCCTGCCGCGCGCTGAGCCCTCCGCCGACCTCTTGGTGAACCTGAGCCGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        757 GACCGCGACCTCCTGCCGCGCGCTGAGCCCTCCGCCGACCTCTTGGTGAACCTGAGCCGC 816

Qy        661 TGCCGACGCCTCATCGTGGTGCTTTCGGACGCCTTCCTGAGCCGGGCCTGGTGCAGCCAC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        817 TGCCGACGCCTCATCGTGGTGCTTTCGGACGCCTTCCTGAGCCGGGCCTGGTGCAGCCAC 876

Qy        721 AGCTTCCGGGAGGGCCTGTGCCGGCTGCTGGAGCTCACCCGCAGACCCATCTTCATCACC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        877 AGCTTCCGGGAGGGCCTGTGCCGGCTGCTGGAGCTCACCCGCAGACCCATCTTCATCACC 936

Qy        781 TTCGAGGGCCAGAGGCGCGACCCCGCGCACCCGGCGCTCCGCCTGCTGCGCCAGCACCGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        937 TTCGAGGGCCAGAGGCGCGACCCCGCGCACCCGGCGCTCCGCCTGCTGCGCCAGCACCGC 996

Qy        841 CACCTGGTGACCTTGCTGCTCTGGAGGCCCGGCTCCGTGACTCCTTCCTCCGATTTTTGG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        997 CACCTGGTGACCTTGCTGCTCTGGAGGCCCGGCTCCGTGACTCCTTCCTCCGATTTTTGG 1056

Qy        901 AAAGAAGTGCAGCTGGCGCTGCCGCGGAAGGTGCGGTACAGGCCGGTGGAAGGAGACCCC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1057 AAAGAAGTGCAGCTGGCGCTGCCGCGGAAGGTGCGGTACAGGCCGGTGGAAGGAGACCCC 1116

Qy        961 CAGACGCAGCTGCAGGACGACAAGGACCCCATGCTGATTCTTCGAGGCCGAGTCCCTGAG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1117 CAGACGCAGCTGCAGGACGACAAGGACCCCATGCTGATTCTTCGAGGCCGAGTCCCTGAG 1176

Qy       1021 GGCCGGGCCCTGGACTCAGAGGTGGACCCGGACCCTGAGGGCGACCTGGGTGTCCGGGGG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1177 GGCCGGGCCCTGGACTCAGAGGTGGACCCGGACCCTGAGGGCGACCTGGGTGTCCGGGGG 1236

Qy       1081 CCTGTTTTTGGAGAGCCATCAGCTCCACCGCACACCAGTGGGGTCTCGCTGGGAGAGAGC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1237 CCTGTTTTTGGAGAGCCATCAGCTCCACCGCACACCAGTGGGGTCTCGCTGGGAGAGAGC 1296

Qy       1141 CGGAGCAGCGAAGTGGACGTCTCGGATCTCGGCTCGCGAAACTACAGTGCCCGCACAGAC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1297 CGGAGCAGCGAAGTGGACGTCTCGGATCTCGGCTCGCGAAACTACAGTGCCCGCACAGAC 1356

Qy       1201 TTCTACTGCCTGGTGTCCAAGGATGATATGTAG 1233
              |||||||||||||||||||||||||||||||||
Db       1357 TTCTACTGCCTGGTGTCCAAGGATGATATGTAG 1389


RESULT 4
AK172830
LOCUS       AK172830               1657 bp    mRNA    linear   PRI 07-MAY-2004
DEFINITION  Homo sapiens cDNA FLJ23991 fis, clone HRC08504, highly similar to
            Homo sapiens single Ig IL-1R-related molecule (SIGIRR).
ACCESSION   AK172830
VERSION     AK172830.1  GI:47077852
KEYWORDS    oligo capping; fis (full insert sequence).
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Watanabe,K., Kumagai,A., Itakura,S., Yamazaki,M., Tashiro,H.,
            Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T.,
            Nakamura,Y., Isogai,T. and Sugano,S.
  TITLE     NEDO human cDNA sequencing project
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1657)
  AUTHORS   Sugano,S. and Suzuki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-APR-2004) Sumio Sugano, Institute of Medical Science,
            University of Tokyo, Laboratory of Genome Structure, Human Genome
            Center; Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan
            (E-mail:flcdna@mail.ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
            Fax:81-3-5449-5416)
COMMENT     NEDO human cDNA sequencing project supported by Ministry of
            Economy, Trade and Industry of Japan; cDNA full insert sequencing:
            Research Association for Biotechnology; cDNA library construction,
            5'- & 3'-end one pass sequencing: Departent of Virology and Human
            Genome Center, Institute of Medical Science, University of Tokyo
            (partly supported by Science and Technology Agency).
FEATURES             Location/Qualifiers
     source          1. .1657
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="HRC08504"
                     /cell_type="primary human renal epithelial cells"
                     /clone_lib="HRC"
                     /note="cloning vector pME18SFL3"
     CDS             227. .1459
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="BAD18795.1"
                     /db_xref="GI:47077853"
                     /translation="MPGVCDRAPDFLSPSEDQVLRPALGSSVALNCTAWVVSGPHCSL
                     PSVQWLKDGLPLGIGGHYSLHEYSWVKANLSEVLVSSVLGVNVTSTEVYGAFTCSIQN
                     ISFSSFTLQRAGPTSHVAAVLASLLVLLALLLAALLYVKCRLNVLLWYQDAYGEVEIN
                     DGKLYDAYVSYSDCPEDRKFVNFILKPQLERRRGYKLFLDDRDLLPRAEPSADLLVNL
                     SRCRRLIVVLSDAFLSRAWCSHSFREGLCRLLELTRRPIFITFEGQRRDPAHPALRLL
                     RQHRHLVTLLLWRPGSVTPSSDFWKEVQLALPRKVRYRPVEGDPQTQLQDDKDPMLIL
                     RGRVPEGRALDSEVDPDPEGDLGVRGPVFGEPSAPPHTSGVSLGESRSSEVDVSDLGS
                     RNYSARTDFYCLVSKDDM"
ORIGIN

  Query Match           100.0%;  Score 1233;  DB 9;  Length 1657;
  Best Local Similarity  100.0%;  Pred. No. 3.1e-182;
  Matches 1233;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGCCAGGTGTCTGTGATAGGGCCCCTGACTTCCTCTCCCCGTCTGAAGACCAGGTGCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 ATGCCAGGTGTCTGTGATAGGGCCCCTGACTTCCTCTCCCCGTCTGAAGACCAGGTGCTG 286

Qy         61 AGGCCTGCCTTGGGCAGCTCAGTGGCTCTGAACTGCACGGCTTGGGTAGTCTCTGGGCCC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 AGGCCTGCCTTGGGCAGCTCAGTGGCTCTGAACTGCACGGCTTGGGTAGTCTCTGGGCCC 346

Qy        121 CACTGCTCCCTGCCTTCAGTCCAGTGGCTGAAAGACGGGCTTCCATTGGGAATTGGGGGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        347 CACTGCTCCCTGCCTTCAGTCCAGTGGCTGAAAGACGGGCTTCCATTGGGAATTGGGGGC 406

Qy        181 CACTACAGCCTCCACGAGTACTCCTGGGTCAAGGCCAACCTGTCAGAGGTGCTTGTGTCC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        407 CACTACAGCCTCCACGAGTACTCCTGGGTCAAGGCCAACCTGTCAGAGGTGCTTGTGTCC 466

Qy        241 AGTGTCCTGGGGGTCAACGTGACCAGCACTGAAGTCTATGGGGCCTTCACCTGCTCCATC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        467 AGTGTCCTGGGGGTCAACGTGACCAGCACTGAAGTCTATGGGGCCTTCACCTGCTCCATC 526

Qy        301 CAGAACATCAGCTTCTCCTCCTTCACTCTTCAGAGAGCTGGCCCTACAAGCCACGTGGCT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        527 CAGAACATCAGCTTCTCCTCCTTCACTCTTCAGAGAGCTGGCCCTACAAGCCACGTGGCT 586

Qy        361 GCGGTGCTGGCCTCCCTCCTGGTCCTGCTGGCCCTGCTGCTGGCCGCCCTGCTCTATGTC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        587 GCGGTGCTGGCCTCCCTCCTGGTCCTGCTGGCCCTGCTGCTGGCCGCCCTGCTCTATGTC 646

Qy        421 AAGTGCCGTCTCAACGTGCTGCTCTGGTACCAGGACGCGTATGGGGAGGTGGAGATAAAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        647 AAGTGCCGTCTCAACGTGCTGCTCTGGTACCAGGACGCGTATGGGGAGGTGGAGATAAAC 706

Qy        481 GACGGGAAGCTCTACGACGCCTACGTCTCCTACAGCGACTGCCCCGAGGACCGCAAGTTC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        707 GACGGGAAGCTCTACGACGCCTACGTCTCCTACAGCGACTGCCCCGAGGACCGCAAGTTC 766

Qy        541 GTGAACTTCATCCTAAAGCCGCAGCTGGAGCGGCGTCGGGGCTACAAGCTCTTCCTGGAC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        767 GTGAACTTCATCCTAAAGCCGCAGCTGGAGCGGCGTCGGGGCTACAAGCTCTTCCTGGAC 826

Qy        601 GACCGCGACCTCCTGCCGCGCGCTGAGCCCTCCGCCGACCTCTTGGTGAACCTGAGCCGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        827 GACCGCGACCTCCTGCCGCGCGCTGAGCCCTCCGCCGACCTCTTGGTGAACCTGAGCCGC 886

Qy        661 TGCCGACGCCTCATCGTGGTGCTTTCGGACGCCTTCCTGAGCCGGGCCTGGTGCAGCCAC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        887 TGCCGACGCCTCATCGTGGTGCTTTCGGACGCCTTCCTGAGCCGGGCCTGGTGCAGCCAC 946

Qy        721 AGCTTCCGGGAGGGCCTGTGCCGGCTGCTGGAGCTCACCCGCAGACCCATCTTCATCACC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        947 AGCTTCCGGGAGGGCCTGTGCCGGCTGCTGGAGCTCACCCGCAGACCCATCTTCATCACC 1006

Qy        781 TTCGAGGGCCAGAGGCGCGACCCCGCGCACCCGGCGCTCCGCCTGCTGCGCCAGCACCGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1007 TTCGAGGGCCAGAGGCGCGACCCCGCGCACCCGGCGCTCCGCCTGCTGCGCCAGCACCGC 1066

Qy        841 CACCTGGTGACCTTGCTGCTCTGGAGGCCCGGCTCCGTGACTCCTTCCTCCGATTTTTGG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1067 CACCTGGTGACCTTGCTGCTCTGGAGGCCCGGCTCCGTGACTCCTTCCTCCGATTTTTGG 1126
```

```
Qy        901 AAAGAAGTGCAGCTGGCGCTGCCGCGGAAGGTGCGGTACAGGCCGGTGGAAGGAGACCCC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1127 AAAGAAGTGCAGCTGGCGCTGCCGCGGAAGGTGCGGTACAGGCCGGTGGAAGGAGACCCC 1186

Qy        961 CAGACGCAGCTGCAGGACGACAAGGACCCCATGCTGATTCTTCGAGGCCGAGTCCCTGAG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1187 CAGACGCAGCTGCAGGACGACAAGGACCCCATGCTGATTCTTCGAGGCCGAGTCCCTGAG 1246

Qy       1021 GGCCGGGCCCTGGACTCAGAGGTGGACCCGGACCCTGAGGGCGACCTGGGTGTCCGGGGG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1247 GGCCGGGCCCTGGACTCAGAGGTGGACCCGGACCCTGAGGGCGACCTGGGTGTCCGGGGG 1306

Qy       1081 CCTGTTTTTGGAGAGCCATCAGCTCCACCGCACACCAGTGGGGTCTCGCTGGGAGAGAGC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1307 CCTGTTTTTGGAGAGCCATCAGCTCCACCGCACACCAGTGGGGTCTCGCTGGGAGAGAGC 1366

Qy       1141 CGGAGCAGCGAAGTGGACGTCTCGGATCTCGGCTCGCGAAACTACAGTGCCCGCACAGAC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1367 CGGAGCAGCGAAGTGGACGTCTCGGATCTCGGCTCGCGAAACTACAGTGCCCGCACAGAC 1426

Qy       1201 TTCTACTGCCTGGTGTCCAAGGATGATATGTAG 1233
              |||||||||||||||||||||||||||||||||
Db       1427 TTCTACTGCCTGGTGTCCAAGGATGATATGTAG 1459


RESULT 5
AK025099
LOCUS       AK025099                1695 bp    mRNA    linear   PRI 13-SEP-2003
DEFINITION  Homo sapiens cDNA: FLJ21446 fis, clone COL04458.
ACCESSION   AK025099
VERSION     AK025099.1  GI:10437546
KEYWORDS    oligo capping; fis (full insert sequence).
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Kawabata,A., Hikiji,T., Kobatake,N., Inagaki,H., Ikema,Y.,
            Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T.,
            Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
  TITLE     NEDO human cDNA sequencing project
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1695)
  AUTHORS   Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
            Shibahara,T., Tanaka,T. and Nakamura,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science,
            University of Tokyo, Laboratory of Genome Structure Analysis, Human
            Genome Center; Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639,
            Japan (E-mail:flcdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
            Fax:81-3-5449-5416)
COMMENT     NEDO human cDNA sequencing project supported by Ministry of
            International Trade and Industry of Japan; cDNA full insert
            sequencing: Research Association for Biotechnology; cDNA library
            construction, 5'- & 3'-end one pass sequencing: Departent of
            Virology and Human Genome Center, Institute of Medical Science,
            University of Tokyo (partly supported by Science and Technology
            Agency).
FEATURES             Location/Qualifiers
     source          1. .1695
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="COL04458"
                     /tissue_type="colon"
                     /clone_lib="COL"
                     /note="cloning vector pME18SFL3"
     CDS             227. .1459
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="BAB15066.1"
                     /db_xref="GI:10437547"
                     /translation="MPGVCDRAPDFLSPSEDQVLRPALGSSVALNCTAWVVSGPHCSL
                     PSVQWLKDGLPLGIGGHYSLHEYSWVKANLSEVLVSSVLGVNVTSTEVYGAFTCSIQN
                     ISFSSFTLQRAGPTSHVAAVLASLLVLLALLLAALLYVKCRLNVLLWYQDAYGEVEIN
                     DGKLYDAYVSYSDCPEDRKFVNFILKPQLERRRGYKLFLDDRDLLPRAEPSADLLVNL
                     SRCRRLIVVLSDAFLSRAWCSHSFREGLCRLLELTRRPIFITFEGQRRDPAHPALRLL
                     RQHRHLVTLLLWRPGSVTPSSDFWKEVQLALPRKVRYRPVEGDPQTQLQDDKDPMLIL
                     RGRVPEGRALDSEVDPDPEGDLGVRGPVFGEPSAPPHTSGVSLGESRSSEVDVSDLGS
                     RNYSARTDFYCLVSKDDM"
ORIGIN

  Query Match             100.0%;  Score 1233;  DB 9;  Length 1695;
  Best Local Similarity   100.0%;  Pred. No. 3.1e-182;
  Matches 1233;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGCCAGGTGTCTGTGATAGGGCCCCTGACTTCCTCTCCCCGTCTGAAGACCAGGTGCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 ATGCCAGGTGTCTGTGATAGGGCCCCTGACTTCCTCTCCCCGTCTGAAGACCAGGTGCTG 286

Qy         61 AGGCCTGCCTTGGGCAGCTCAGTGGCTCTGAACTGCACGGCTTGGGTAGTCTCTGGGCCC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 AGGCCTGCCTTGGGCAGCTCAGTGGCTCTGAACTGCACGGCTTGGGTAGTCTCTGGGCCC 346

Qy        121 CACTGCTCCCTGCCTTCAGTCCAGTGGCTGAAAGACGGGCTTCCATTGGGAATTGGGGGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        347 CACTGCTCCCTGCCTTCAGTCCAGTGGCTGAAAGACGGGCTTCCATTGGGAATTGGGGGC 406

Qy        181 CACTACAGCCTCCACGAGTACTCCTGGGTCAAGGCCAACCTGTCAGAGGTGCTTGTGTCC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        407 CACTACAGCCTCCACGAGTACTCCTGGGTCAAGGCCAACCTGTCAGAGGTGCTTGTGTCC 466

Qy        241 AGTGTCCTGGGGGTCAACGTGACCAGCACTGAAGTCTATGGGGCCTTCACCTGCTCCATC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        467 AGTGTCCTGGGGGTCAACGTGACCAGCACTGAAGTCTATGGGGCCTTCACCTGCTCCATC 526

Qy        301 CAGAACATCAGCTTCTCCTCCTTCACTCTTCAGAGAGCTGGCCCTACAAGCCACGTGGCT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        527 CAGAACATCAGCTTCTCCTCCTTCACTCTTCAGAGAGCTGGCCCTACAAGCCACGTGGCT 586

Qy        361 GCGGTGCTGGCCTCCCTCCTGGTCCTGCTGGCCCTGCTGCTGGCCGCCCTGCTCTATGTC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        587 GCGGTGCTGGCCTCCCTCCTGGTCCTGCTGGCCCTGCTGCTGGCCGCCCTGCTCTATGTC 646

Qy        421 AAGTGCCGTCTCAACGTGCTGCTCTGGTACCAGGACGCGTATGGGGAGGTGGAGATAAAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        647 AAGTGCCGTCTCAACGTGCTGCTCTGGTACCAGGACGCGTATGGGGAGGTGGAGATAAAC 706

Qy        481 GACGGGAAGCTCTACGACGCCTACGTCTCCTACAGCGACTGCCCCGAGGACCGCAAGTTC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        707 GACGGGAAGCTCTACGACGCCTACGTCTCCTACAGCGACTGCCCCGAGGACCGCAAGTTC 766

Qy        541 GTGAACTTCATCCTAAAGCCGCAGCTGGAGCGGCGTCGGGGCTACAAGCTCTTCCTGGAC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        767 GTGAACTTCATCCTAAAGCCGCAGCTGGAGCGGCGTCGGGGCTACAAGCTCTTCCTGGAC 826

Qy        601 GACCGCGACCTCCTGCCGCGCGCTGAGCCCTCCGCCGACCTCTTGGTGAACCTGAGCCGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        827 GACCGCGACCTCCTGCCGCGCGCTGAGCCCTCCGCCGACCTCTTGGTGAACCTGAGCCGC 886

Qy        661 TGCCGACGCCTCATCGTGGTGCTTTCGGACGCCTTCCTGAGCCGGGCCTGGTGCAGCCAC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        887 TGCCGACGCCTCATCGTGGTGCTTTCGGACGCCTTCCTGAGCCGGGCCTGGTGCAGCCAC 946

Qy        721 AGCTTCCGGGAGGGCCTGTGCCGGCTGCTGGAGCTCACCCGCAGACCCATCTTCATCACC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        947 AGCTTCCGGGAGGGCCTGTGCCGGCTGCTGGAGCTCACCCGCAGACCCATCTTCATCACC 1006

Qy        781 TTCGAGGGCCAGAGGCGCGACCCCGCGCACCCGGCGCTCCGCCTGCTGCGCCAGCACCGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1007 TTCGAGGGCCAGAGGCGCGACCCCGCGCACCCGGCGCTCCGCCTGCTGCGCCAGCACCGC 1066

Qy        841 CACCTGGTGACCTTGCTGCTCTGGAGGCCCGGCTCCGTGACTCCTTCCTCCGATTTTTGG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1067 CACCTGGTGACCTTGCTGCTCTGGAGGCCCGGCTCCGTGACTCCTTCCTCCGATTTTTGG 1126
```

```
Qy        901 AAAGAAGTGCAGCTGGCGCTGCCGCGGAAGGTGCGGTACAGGCCGGTGGAAGGAGACCCC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1127 AAAGAAGTGCAGCTGGCGCTGCCGCGGAAGGTGCGGTACAGGCCGGTGGAAGGAGACCCC 1186

Qy        961 CAGACGCAGCTGCAGGACGACAAGGACCCCATGCTGATTCTTCGAGGCCGAGTCCCTGAG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1187 CAGACGCAGCTGCAGGACGACAAGGACCCCATGCTGATTCTTCGAGGCCGAGTCCCTGAG 1246

Qy       1021 GGCCGGGCCCTGGACTCAGAGGTGGACCCGGACCCTGAGGGCGACCTGGGTGTCCGGGGG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1247 GGCCGGGCCCTGGACTCAGAGGTGGACCCGGACCCTGAGGGCGACCTGGGTGTCCGGGGG 1306

Qy       1081 CCTGTTTTTGGAGAGCCATCAGCTCCACCGCACACCAGTGGGGTCTCGCTGGGAGAGAGC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1307 CCTGTTTTTGGAGAGCCATCAGCTCCACCGCACACCAGTGGGGTCTCGCTGGGAGAGAGC 1366

Qy       1141 CGGAGCAGCGAAGTGGACGTCTCGGATCTCGGCTCGCGAAACTACAGTGCCCGCACAGAC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1367 CGGAGCAGCGAAGTGGACGTCTCGGATCTCGGCTCGCGAAACTACAGTGCCCGCACAGAC 1426

Qy       1201 TTCTACTGCCTGGTGTCCAAGGATGATATGTAG 1233
              ||||||||||||||||||||||||||||||||
Db       1427 TTCTACTGCCTGGTGTCCAAGGATGATATGTAG 1459
```

```
RESULT 6
AX058610
LOCUS       AX058610               1659 bp    DNA     linear   PAT 17-JAN-2001
DEFINITION  Sequence 17 from Patent WO0075321.
ACCESSION   AX058610
VERSION     AX058610.1  GI:12310952
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Shimkets,R.A., Fernandes,E., Herrman,J. and Vernet,C.
  TITLE     Polynucleotides and membrane-bound polypeptides encoded thereby
  JOURNAL   Patent: WO 0075321-A 17 14-DEC-2000;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .1659
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             244. .1476
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC22520.1"
                     /db_xref="GI:12310953"
                     /translation="MPGVCDRAPDFLSPSEDQVLRPALGSSVALNCTAWVVSGPHCSL
                     PSVQWLKDGLPLGIGGHYSLHEYSWVKANLSEVLVSSVLGVNVTSTEVYGAFTCSIQN
                     ISFSSFTLQRAGPTSHVAAVLASLLVLLALLLAALLYVKCRLNVLLWYQDAYGEVEIN
                     DGKLYDAYVSYSDCPEDRKFVNFILKPQLERRRGYKLFLDDRDLLPRAEPSADLLVNL
                     SRCRRLIVVLSDAFLSRAWCSHSFREGLCRLLELTRRPIFITFEGQRRDPAHPALRLL
                     RQHRHLVTLLLWRPGSVTPSSDFWKEVQLALPRKVRYRPVEGDPQTQLQDDKDPMLIL
                     RGRVPEGRALDSEVDPDPEGDLGVRGPVFGEPSAPPHTSGVSLGESRSSEVDVSDLGS
                     RNYSARTDFYCLVSKDDM"
ORIGIN

  Query Match             99.9%;  Score 1231.4;  DB 6;  Length 1659;
  Best Local Similarity   99.9%;  Pred. No. 5.6e-182;
  Matches 1232;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 ATGCCAGGTGTCTGTGATAGGGCCCCTGACTTCCTCTCCCCGTCTGAAGACCAGGTGCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 ATGCCAGGTGTCTGTGATAGGGCCCCTGACTTCCTCTCCCCGTCTGAAGACCAGGTGCTG 303

Qy         61 AGGCCTGCCTTGGGCAGCTCAGTGGCTCTGAACTGCACGGCTTGGGTAGTCTCTGGGCCC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        304 AGGCCTGCCTTGGGCAGCTCAGTGGCTCTGAACTGCACGGCTTGGGTAGTCTCTGGGCCC 363
```

```
Qy        121 CACTGCTCCCTGCCTTCAGTCCAGTGGCTGAAAGACGGGCTTCCATTGGGAATTGGGGGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        364 CACTGCTCCCTGCCTTCAGTCCAGTGGCTGAAAGACGGGCTTCCATTGGGAATTGGGGGC 423

Qy        181 CACTACAGCCTCCACGAGTACTCCTGGGTCAAGGCCAACCTGTCAGAGGTGCTTGTGTCC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        424 CACTACAGCCTCCACGAGTACTCCTGGGTCAAGGCCAACCTGTCAGAGGTGCTTGTGTCC 483

Qy        241 AGTGTCCTGGGGGTCAACGTGACCAGCACTGAAGTCTATGGGGCCTTCACCTGCTCCATC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        484 AGTGTCCTGGGGGTCAACGTGACCAGCACTGAAGTCTATGGGGCCTTCACCTGCTCCATC 543

Qy        301 CAGAACATCAGCTTCTCCTCCTTCACTCTTCAGAGAGCTGGCCCTACAAGCCACGTGGCT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        544 CAGAACATCAGCTTCTCCTCCTTCACTCTTCAGAGAGCTGGCCCTACAAGCCACGTGGCT 603

Qy        361 GCGGTGCTGGCCTCCCTCCTGGTCCTGCTGGCCCTGCTGCTGGCCGCCCTGCTCTATGTC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        604 GCGGTGCTGGCCTCCCTCCTGGTCCTGCTGGCCCTGCTGCTGGCCGCCCTGCTCTATGTC 663

Qy        421 AAGTGCCGTCTCAACGTGCTGCTCTGGTACCAGGACGCGTATGGGGAGGTGGAGATAAAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        664 AAGTGCCGTCTCAACGTGCTGCTCTGGTACCAGGACGCGTATGGGGAGGTGGAGATAAAC 723

Qy        481 GACGGGAAGCTCTACGACGCCTACGTCTCCTACAGCGACTGCCCCGAGGACCGCAAGTTC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        724 GACGGGAAGCTCTACGACGCCTACGTCTCCTACAGCGACTGCCCCGAGGACCGCAAGTTC 783

Qy        541 GTGAACTTCATCCTAAAGCCGCAGCTGGAGCGGCGTCGGGGCTACAAGCTCTTCCTGGAC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        784 GTGAACTTCATCCTAAAGCCGCAGCTGGAGCGGCGTCGGGGCTACAAGCTCTTCCTGGAC 843

Qy        601 GACCGCGACCTCCTGCCGCGCGCTGAGCCCTCCGCCGACCTCTTGGTGAACCTGAGCCGC 660
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        844 GACCGCGACCTCCTGCCCCGCGCTGAGCCCTCCGCCGACCTCTTGGTGAACCTGAGCCGC 903

Qy        661 TGCCGACGCCTCATCGTGGTGCTTTCGGACGCCTTCCTGAGCCGGGCCTGGTGCAGCCAC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        904 TGCCGACGCCTCATCGTGGTGCTTTCGGACGCCTTCCTGAGCCGGGCCTGGTGCAGCCAC 963

Qy        721 AGCTTCCGGGAGGGCCTGTGCCGGCTGCTGGAGCTCACCCGCAGACCCATCTTCATCACC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        964 AGCTTCCGGGAGGGCCTGTGCCGGCTGCTGGAGCTCACCCGCAGACCCATCTTCATCACC 1023

Qy        781 TTCGAGGGCCAGAGGCGCGACCCCGCGCACCCGGCGCTCCGCCTGCTGCGCCAGCACCGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1024 TTCGAGGGCCAGAGGCGCGACCCCGCGCACCCGGCGCTCCGCCTGCTGCGCCAGCACCGC 1083

Qy        841 CACCTGGTGACCTTGCTGCTCTGGAGGCCCGGCTCCGTGACTCCTTCCTCCGATTTTTGG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1084 CACCTGGTGACCTTGCTGCTCTGGAGGCCCGGCTCCGTGACTCCTTCCTCCGATTTTTGG 1143

Qy        901 AAAGAAGTGCAGCTGGCGCTGCCGCGGAAGGTGCGGTACAGGCCGGTGGAAGGAGACCCC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1144 AAAGAAGTGCAGCTGGCGCTGCCGCGGAAGGTGCGGTACAGGCCGGTGGAAGGAGACCCC 1203

Qy        961 CAGACGCAGCTGCAGGACGACAAGGACCCCATGCTGATTCTTCGAGGCCGAGTCCCTGAG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1204 CAGACGCAGCTGCAGGACGACAAGGACCCCATGCTGATTCTTCGAGGCCGAGTCCCTGAG 1263

Qy       1021 GGCCGGGCCCTGGACTCAGAGGTGGACCCGGACCCTGAGGGCGACCTGGGTGTCCGGGGG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1264 GGCCGGGCCCTGGACTCAGAGGTGGACCCGGACCCTGAGGGCGACCTGGGTGTCCGGGGG 1323

Qy       1081 CCTGTTTTTGGAGAGCCATCAGCTCCACCGCACACCAGTGGGGTCTCGCTGGGAGAGAGC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1324 CCTGTTTTTGGAGAGCCATCAGCTCCACCGCACACCAGTGGGGTCTCGCTGGGAGAGAGC 1383

Qy       1141 CGGAGCAGCGAAGTGGACGTCTCGGATCTCGGCTCGCGAAACTACAGTGCCCGCACAGAC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1384 CGGAGCAGCGAAGTGGACGTCTCGGATCTCGGCTCGCGAAACTACAGTGCCCGCACAGAC 1443
```

```
Qy       1201 TTCTACTGCCTGGTGTCCAAGGATGATATGTAG 1233
              |||||||||||||||||||||||||||||||||
Db       1444 TTCTACTGCCTGGTGTCCAAGGATGATATGTAG 1476


RESULT 7
CR457338
LOCUS       CR457338               1233 bp    mRNA    linear   PRI 03-JUN-2004
DEFINITION  Homo sapiens full open reading frame cDNA clone RZPDo834H0613D for
            gene SIGIRR, single Ig IL-1R-related molecule; complete cds, incl.
            stopcodon.
ACCESSION   CR457338
VERSION     CR457338.1  GI:48146792
KEYWORDS    Full ORF shuttle clone, Gateway(TM), complete cds.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1233)
  AUTHORS   Ebert,L., Schick,M., Neubert,P., Schatten,R., Henze,S. and Korn,B.
  TITLE     Cloning of human full open reading frames in Gateway(TM) system
            entry vector (pDONR201)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1233)
  AUTHORS   Ebert,L., Schick,M., Neubert,P., Schatten,R., Henze,S. and Korn,B.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-JUN-2004) RZPD Deutsches Ressourcenzentrum fuer
            Genomforschung GmbH, Im Neuenheimer Feld 580, D-69120 Heidelberg,
            Germany
COMMENT     RZPD; RZPDo834H0613D, ORFNo 2417
            www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=RZPDo834H0613D RZPDLIB;
            Human Full ORF Clones Gateway(TM) - RZPD (kan-resist.) RZPD LIB No.
            834
            www.rzpd.de/cgi-bin/products/showLib.pl.cgi/response?libNo=834
            www.rzpd.de/products/orfclones/
            Contact: Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH,
            Heubnerweg 6, D-14059 Berlin, Germany
            Tel: +49 30 32639 100
            Fax: +49 30 32639 111
            www.rzpd.de
            This clone is available from RZPD;
            contact RZPD (customer.service@rzpd.de) for further information.
            This CDS clone is a part of a collection of human full length
            expression clones generated by RZPD.
            This CDS has been cloned incl. stopcodon.
            The CDS has been inserted into pDONR201 via a BP Clonase(TM)
            reaction. Additional sequence has been added in front of the start
            codon: att. .AAAAAA GCA GGC (ATG).
            The last base of the last coding triplett has been changed to T,
            which might lead to an amino acid change at the C terminus of the
            polypeptide.
            The stop codon has been set to TAA followed by
            TTAACCCAGCTTTCTT. .att. Compared to the reference sequence NM_021805
            we found amino acid exchange(s) at position (first base of changed
            triplet): 1228(met->ile)
            Clone distribution: http://www.rzpd.de/products/orfclones/.
FEATURES             Location/Qualifiers
     source          1. .1233
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="RZPDo834H0613D"
                     /clone_lib="Human Full ORF Clones Gateway(TM) - RZPD"
                     /lab_host="DH10B"
                     /note="Vector: pDONR201, Site_1: attP1; Site_2: attP2"
     gene            1. .1233
                     /gene="SIGIRR"
     CDS             1. .1233
                     /gene="SIGIRR"
                     /codon_start=1
                     /protein_id="CAG33619.1"
                     /db_xref="GI:48146793"
                     /translation="MPGVCDRAPDFLSPSEDQVLRPALGSSVALNCTAWVVSGPHCSL
                     PSVQWLKDGLPLGIGGHYSLHEYSWVKANLSEVLVSSVLGVNVTSTEVYGAFTCSIQN
                     ISFSSFTLQRAGPTSHVAAVLASLLVLLALLLAALLYVKCRLNVLLWYQDAYGEVEIN
                     DGKLYDAYVSYSDCPEDRKFVNFILKPQLERRRGYKLFLDDRDLLPRAEPSADLLVNL
                     SRCRRLIVVLSDAFLSRAWCSHSFREGLCRLLELTRRPIFITFEGQRRDPAHPALRLL
                     RQHRHLVTLLLWRPGSVTPSSDFWKEVQLALPRKVRYRPVEGDPQTQLQDDKDPMLIL
                     RGRVPEGRALDSEVDPDPEGDLGVRGPVFGEPSAPPHTSGVSLGESRSSEVDVSDLGS
                     RNYSARTDFYCLVSKDDI"
ORIGIN

  Query Match             99.8%;  Score 1230.4;  DB 9;  Length 1233;
  Best Local Similarity   99.9%;  Pred. No. 8.6e-182;
  Matches 1231;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 ATGCCAGGTGTCTGTGATAGGGCCCCTGACTTCCTCTCCCCGTCTGAAGACCAGGTGCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCCAGGTGTCTGTGATAGGGCCCCTGACTTCCTCTCCCCGTCTGAAGACCAGGTGCTG 60

Qy         61 AGGCCTGCCTTGGGCAGCTCAGTGGCTCTGAACTGCACGGCTTGGGTAGTCTCTGGGCCC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGGCCTGCCTTGGGCAGCTCAGTGGCTCTGAACTGCACGGCTTGGGTAGTCTCTGGGCCC 120

Qy        121 CACTGCTCCCTGCCTTCAGTCCAGTGGCTGAAAGACGGGCTTCCATTGGGAATTGGGGGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CACTGCTCCCTGCCTTCAGTCCAGTGGCTGAAAGACGGGCTTCCATTGGGAATTGGGGGC 180

Qy        181 CACTACAGCCTCCACGAGTACTCCTGGGTCAAGGCCAACCTGTCAGAGGTGCTTGTGTCC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CACTACAGCCTCCACGAGTACTCCTGGGTCAAGGCCAACCTGTCAGAGGTGCTTGTGTCC 240

Qy        241 AGTGTCCTGGGGGTCAACGTGACCAGCACTGAAGTCTATGGGGCCTTCACCTGCTCCATC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGTGTCCTGGGGGTCAACGTGACCAGCACTGAAGTCTATGGGGCCTTCACCTGCTCCATC 300

Qy        301 CAGAACATCAGCTTCTCCTCCTTCACTCTTCAGAGAGCTGGCCCTACAAGCCACGTGGCT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CAGAACATCAGCTTCTCCTCCTTCACTCTTCAGAGAGCTGGCCCTACAAGCCACGTGGCT 360

Qy        361 GCGGTGCTGGCCTCCCTCCTGGTCCTGCTGGCCCTGCTGCTGGCCGCCCTGCTCTATGTC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCGGTGCTGGCCTCCCTCCTGGTCCTGCTGGCCCTGCTGCTGGCCGCCCTGCTCTATGTC 420

Qy        421 AAGTGCCGTCTCAACGTGCTGCTCTGGTACCAGGACGCGTATGGGGAGGTGGAGATAAAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AAGTGCCGTCTCAACGTGCTGCTCTGGTACCAGGACGCGTATGGGGAGGTGGAGATAAAC 480

Qy        481 GACGGGAAGCTCTACGACGCCTACGTCTCCTACAGCGACTGCCCCGAGGACCGCAAGTTC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GACGGGAAGCTCTACGACGCCTACGTCTCCTACAGCGACTGCCCCGAGGACCGCAAGTTC 540

Qy        541 GTGAACTTCATCCTAAAGCCGCAGCTGGAGCGGCGTCGGGGCTACAAGCTCTTCCTGGAC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GTGAACTTCATCCTAAAGCCGCAGCTGGAGCGGCGTCGGGGCTACAAGCTCTTCCTGGAC 600

Qy        601 GACCGCGACCTCCTGCCGCGCGCTGAGCCCTCCGCCGACCTCTTGGTGAACCTGAGCCGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GACCGCGACCTCCTGCCGCGCGCTGAGCCCTCCGCCGACCTCTTGGTGAACCTGAGCCGC 660

Qy        661 TGCCGACGCCTCATCGTGGTGCTTTCGGACGCCTTCCTGAGCCGGGCCTGGTGCAGCCAC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGCCGACGCCTCATCGTGGTGCTTTCGGACGCCTTCCTGAGCCGGGCCTGGTGCAGCCAC 720

Qy        721 AGCTTCCGGGAGGGCCTGTGCCGGCTGCTGGAGCTCACCCGCAGACCCATCTTCATCACC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AGCTTCCGGGAGGGCCTGTGCCGGCTGCTGGAGCTCACCCGCAGACCCATCTTCATCACC 780

Qy        781 TTCGAGGGCCAGAGGCGCGACCCCGCGCACCCGGCGCTCCGCCTGCTGCGCCAGCACCGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TTCGAGGGCCAGAGGCGCGACCCCGCGCACCCGGCGCTCCGCCTGCTGCGCCAGCACCGC 840

Qy        841 CACCTGGTGACCTTGCTGCTCTGGAGGCCCGGCTCCGTGACTCCTTCCTCCGATTTTTGG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CACCTGGTGACCTTGCTGCTCTGGAGGCCCGGCTCCGTGACTCCTTCCTCCGATTTTTGG 900
```

```
Qy        901 AAAGAAGTGCAGCTGGCGCTGCCGCGGAAGGTGCGGTACAGGCCGGTGGAAGGAGACCCC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AAAGAAGTGCAGCTGGCGCTGCCGCGGAAGGTGCGGTACAGGCCGGTGGAAGGAGACCCC 960

Qy        961 CAGACGCAGCTGCAGGACGACAAGGACCCCATGCTGATTCTTCGAGGCCGAGTCCCTGAG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CAGACGCAGCTGCAGGACGACAAGGACCCCATGCTGATTCTTCGAGGCCGAGTCCCTGAG 1020

Qy       1021 GGCCGGGCCCTGGACTCAGAGGTGGACCCGGACCCTGAGGGCGACCTGGGTGTCCGGGGG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GGCCGGGCCCTGGACTCAGAGGTGGACCCGGACCCTGAGGGCGACCTGGGTGTCCGGGGG 1080

Qy       1081 CCTGTTTTTGGAGAGCCATCAGCTCCACCGCACACCAGTGGGGTCTCGCTGGGAGAGAGC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CCTGTTTTTGGAGAGCCATCAGCTCCACCGCACACCAGTGGGGTCTCGCTGGGAGAGAGC 1140

Qy       1141 CGGAGCAGCGAAGTGGACGTCTCGGATCTCGGCTCGCGAAACTACAGTGCCCGCACAGAC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CGGAGCAGCGAAGTGGACGTCTCGGATCTCGGCTCGCGAAACTACAGTGCCCGCACAGAC 1200

Qy       1201 TTCTACTGCCTGGTGTCCAAGGATGATATGTA 1232
              |||||||||||||||||||||||||||||| ||
Db       1201 TTCTACTGCCTGGTGTCCAAGGATGATATTTA 1232
```

```
RESULT 8
AX201368
LOCUS       AX201368                1649 bp    DNA     linear   PAT 30-AUG-2001
DEFINITION  Sequence 47 from Patent WO0153486.
ACCESSION   AX201368
VERSION     AX201368.1  GI:15391178
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Ashkenazi,A.J., Goddard,A., Godowski,P.J., Gurney,A.L.,
            Hillan,K.J., Marsters,S.A., Pan,J., Pitti,R.M., Roy,M.A., Smith,V.,
            Stone,D.M., Watanabe,C.K. and Wood,W.I.
  TITLE     Compositions and methods for the treatment of tumour
  JOURNAL   Patent: WO 0153486-A 47 26-JUL-2001;
            Genentech, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .1649
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN
```

```
  Query Match             86.9%;  Score 1071.8;  DB 6;  Length 1649;
  Best Local Similarity   99.8%;  Pred. No. 3.6e-157;
  Matches 1073;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 ATGCCAGGTGTCTGTGATAGGGCCCCTGACTTCCTCTCCCCGTCTGAAGACCAGGTGCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        104 ATGCCAGGTGTCTGTGATAGGGCCCCTGACTTCCTCTCCCCGTCTGAAGACCAGGTGCTG 163

Qy         61 AGGCCTGCCTTGGGCAGCTCAGTGGCTCTGAACTGCACGGCTTGGGTAGTCTCTGGGCCC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        164 AGGCCTGCCTTGGGCAGCTCAGTGGCTCTGAACTGCACGGCTTGGGTAGTCTCTGGGCCC 223

Qy        121 CACTGCTCCCTGCCTTCAGTCCAGTGGCTGAAAGACGGGCTTCCATTGGGAATTGGGGGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        224 CACTGCTCCCTGCCTTCAGTCCAGTGGCTGAAAGACGGGCTTCCATTGGGAATTGGGGGC 283

Qy        181 CACTACAGCCTCCACGAGTACTCCTGGGTCAAGGCCAACCTGTCAGAGGTGCTTGTGTCC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        284 CACTACAGCCTCCACGAGTACTCCTGGGTCAAGGCCAACCTGTCAGAGGTGCTTGTGTCC 343

Qy        241 AGTGTCCTGGGGGTCAACGTGACCAGCACTGAAGTCTATGGGGCCTTCACCTGCTCCATC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        344 AGTGTCCTGGGGGTCAACGTGACCAGCACTGAAGTCTATGGGGCCTTCACCTGCTCCATC 403

Qy        301 CAGAACATCAGCTTCTCCTCCTTCACTCTTCAGAGAGCTGGCCCTACAAGCCACGTGGCT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        404 CAGAACATCAGCTTCTCCTCCTTCACTCTTCAGAGAGCTGGCCCTACAAGCCACGTGGCT 463

Qy        361 GCGGTGCTGGCCTCCCTCCTGGTCCTGCTGGCCCTGCTGCTGGCCGCCCTGCTCTATGTC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        464 GCGGTGCTGGCCTCCCTCCTGGTCCTGCTGGCCCTGCTGCTGGCCGCCCTGCTCTATGTC 523

Qy        421 AAGTGCCGTCTCAACGTGCTGCTCTGGTACCAGGACGCGTATGGGGAGGTGGAGATAAAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        524 AAGTGCCGTCTCAACGTGCTGCTCTGGTACCAGGACGCGTATGGGGAGGTGGAGATAAAC 583

Qy        481 GACGGGAAGCTCTACGACGCCTACGTCTCCTACAGCGACTGCCCCGAGGACCGCAAGTTC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        584 GACGGGAAGCTCTACGACGCCTACGTCTCCTACAGCGACTGCCCCGAGGACCGCAAGTTC 643

Qy        541 GTGAACTTCATCCTAAAGCCGCAGCTGGAGCGGCGTCGGGGCTACAAGCTCTTCCTGGAC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        644 GTGAACTTCATCCTAAAGCCGCAGCTGGAGCGGCGTCGGGGCTACAAGCTCTTCCTGGAC 703

Qy        601 GACCGCGACCTCCTGCCGCGCGCGCTGAGCCCTCCGCCGACCTCTTGGTGAACCTGAGCCGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        704 GACCGCGACCTCCTGCCGCGCGCGCTGAGCCCTCCGCCGACCTCTTGGTGAACCTGAGCCGC 763

Qy        661 TGCCGACGCCTCATCGTGGTGCTTTCGGACGCCTTCCTGAGCCGGGCCTGGTGCAGCCAC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        764 TGCCGACGCCTCATCGTGGTGCTTTCGGACGCCTTCCTGAGCCGGGCCTGGTGCAGCCAC 823

Qy        721 AGCTTCCGGGAGGGCCTGTGCCGGCTGCTGGAGCTCACCCGCAGACCCATCTTCATCACC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        824 AGCTTCCGGGAGGGCCTGTGCCGGCTGCTGGAGCTCACCCGCAGACCCATCTTCATCACC 883

Qy        781 TTCGAGGGCCAGAGGCGCGACCCCGCGCACCCGGCGCTCCGCCTGCTGCGCCAGCACCGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        884 TTCGAGGGCCAGAGGCGCGACCCCGCGCACCCGGCGCTCCGCCTGCTGCGCCAGCACCGC 943

Qy        841 CACCTGGTGACCTTGCTGCTCTGGAGGCCCGGCTCCGTGACTCCTTCCTCCGATTTTTGG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        944 CACCTGGTGACCTTGCTGCTCTGGAGGCCCGGCTCCGTGACTCCTTCCTCCGATTTTTGG 1003

Qy        901 AAAGAAGTGCAGCTGGCGCTGCCGCGGAAGGTGCGGTACAGGCCGGTGGAAGGAGACCCC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1004 AAAGAAGTGCAGCTGGCGCTGCCGCGGAAGGTGCGGTACAGGCCGGTGGAAGGAGACCCC 1063

Qy        961 CAGACGCAGCTGCAGGACGACAAGGACCCCATGCTGATTCTTCGAGGCCGAGTCCCTGAG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1064 CAGACGCAGCTGCAGGACGACAAGGACCCCATGCTGATTCTTCGAGGCCGAGTCCCTGAG 1123

Qy       1021 GGCCGGGCCCTGGACTCAGAGGTGGACCCGGACCCTGAGGGCGACCTGGGTGTCC 1075
              ||||||||||||||||||||||||||||||||||||||||||||||||||||| | |
Db       1124 GGCCGGGCCCTGGACTCAGAGGTGGACCCGGACCCTGAGGGCGACCTGGGTATGC 1178
```

```
RESULT 9
AY358342
LOCUS       AY358342                1649 bp    mRNA    linear   PRI 03-OCT-2003
DEFINITION  Homo sapiens clone DNA38649 SIGIRR (UNQ301) mRNA, complete cds.
ACCESSION   AY358342
VERSION     AY358342.1  GI:37181808
KEYWORDS    FLI_CDNA.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1649)
  AUTHORS   Clark,H.F., Gurney,A.L., Abaya,E., Baker,K., Baldwin,D., Brush,J.,
            Chen,J., Chow,B., Chui,C., Crowley,C., Currell,B., Deuel,B.,
            Dowd,P., Eaton,D., Foster,J., Grimaldi,C., Gu,Q., Hass,P.E.,
            Heldens,S., Huang,A., Kim,H.S., Klimowski,L., Jin,Y., Johnson,S.,
            Lee,J., Lewis,L., Liao,D., Mark,M., Robbie,E., Sanchez,C.,
            Schoenfeld,J., Seshagiri,S., Simmons,L., Singh,J., Smith,V.,
```

```
            Stinson,J., Vagts,A., Vandlen,R., Watanabe,C., Wieand,D., Woods,K.,
            Xie,M.H., Yansura,D., Yi,S., Yu,G., Yuan,J., Zhang,M., Zhang,Z.,
            Goddard,A., Wood,W.I. and Godowski,P.
  TITLE     The Secreted Protein Discovery Initiative (SPDI), a Large-Scale
            Effort to Identify Novel Human Secreted and Transmembrane Proteins:
            A Bioinformatics Assessment
  JOURNAL   Genome Res. 13 (10), 2265-2270 (2003)
   PUBMED   12975309
REFERENCE   2  (bases 1 to 1649)
  AUTHORS   Clark,H.F.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-AUG-2003) Department of Bioinformatics, Genentech,
            Inc., 1 DNA Way, South San Francisco, CA 94080, USA
FEATURES             Location/Qualifiers
     source          1. .1649
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="DNA38649"
     gene            1. .1649
                     /locus_tag="UNQ301"
     CDS             104. .1618
                     /locus_tag="UNQ301"
                     /note="PRO342"
                     /codon_start=1
                     /product="SIGIRR"
                     /protein_id="AAQ88708.1"
                     /db_xref="GI:37181809"
                     /translation="MPGVCDRAPDFLSPSEDQVLRPALGSSVALNCTAWVVSGPHCSL
                     PSVQWLKDGLPLGIGGHYSLHEYSWVKANLSEVLVSSVLGVNVTSTEVYGAFTCSIQN
                     ISFSSFTLQRAGPTSHVAAVLASLLVLLALLLAALLYVKCRLNVLLWYQDAYGEVEIN
                     DGKLYDAYVSYSDCPEDRKFVNFILKPQLERRRGYKLFLDDRDLLPRAEPSADLLVNL
                     SRCRRLIVVLSDAFLSRAWCSHSFREGLCRLLELTRRPIFITFEGQRRDPAHPALRLL
                     RQHRHLVTLLLWRPGSVTPSSDFWKEVQLALPRKVRYRPVEGDPQTQLQDDKDPMLIL
                     RGRVPEGRALDSEVDPDPEGDLGMPAQPHSPTGEAQHRAEWGQAQGTGPGGALGVEDS
                     SRHREPLHGLCPGGARPSVCLGTSWASQAITAGGEQGQPLAVGLGQGCGWPPQASRSP
                     HPRCPGACFWRAISSTAHQWGLAGREPEQRSGRLGSRLAKLQCPHRLLLPGVQG"
ORIGIN

  Query Match             86.9%;  Score 1071.8;  DB 9;  Length 1649;
  Best Local Similarity   99.8%;  Pred. No. 3.6e-157;
  Matches 1073;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 ATGCCAGGTGTCTGTGATAGGGCCCCTGACTTCCTCTCCCCGTCTGAAGACCAGGTGCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        104 ATGCCAGGTGTCTGTGATAGGGCCCCTGACTTCCTCTCCCCGTCTGAAGACCAGGTGCTG 163

Qy         61 AGGCCTGCCTTGGGCAGCTCAGTGGCTCTGAACTGCACGGCTTGGGTAGTCTCTGGGCCC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        164 AGGCCTGCCTTGGGCAGCTCAGTGGCTCTGAACTGCACGGCTTGGGTAGTCTCTGGGCCC 223

Qy        121 CACTGCTCCCTGCCTTCAGTCCAGTGGCTGAAAGACGGGCTTCCATTGGGAATTGGGGGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        224 CACTGCTCCCTGCCTTCAGTCCAGTGGCTGAAAGACGGGCTTCCATTGGGAATTGGGGGC 283

Qy        181 CACTACAGCCTCCACGAGTACTCCTGGGTCAAGGCCAACCTGTCAGAGGTGCTTGTGTCC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        284 CACTACAGCCTCCACGAGTACTCCTGGGTCAAGGCCAACCTGTCAGAGGTGCTTGTGTCC 343

Qy        241 AGTGTCCTGGGGGTCAACGTGACCAGCACTGAAGTCTATGGGGCCTTCACCTGCTCCATC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        344 AGTGTCCTGGGGGTCAACGTGACCAGCACTGAAGTCTATGGGGCCTTCACCTGCTCCATC 403

Qy        301 CAGAACATCAGCTTCTCCTCCTTCACTCTTCAGAGAGCTGGCCCTACAAGCCACGTGGCT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        404 CAGAACATCAGCTTCTCCTCCTTCACTCTTCAGAGAGCTGGCCCTACAAGCCACGTGGCT 463

Qy        361 GCGGTGCTGGCCTCCCTCCTGGTCCTGCTGGCCCTGCTGCTGGCCGCCCTGCTCTATGTC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        464 GCGGTGCTGGCCTCCCTCCTGGTCCTGCTGGCCCTGCTGCTGGCCGCCCTGCTCTATGTC 523

Qy        421 AAGTGCCGTCTCAACGTGCTGCTCTGGTACCAGGACGCGTATGGGGAGGTGGAGATAAAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        524 AAGTGCCGTCTCAACGTGCTGCTCTGGTACCAGGACGCGTATGGGGAGGTGGAGATAAAC 583

Qy        481 GACGGGAAGCTCTACGACGCCTACGTCTCCTACAGCGACTGCCCCGAGGACCGCAAGTTC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        584 GACGGGAAGCTCTACGACGCCTACGTCTCCTACAGCGACTGCCCCGAGGACCGCAAGTTC 643

Qy        541 GTGAACTTCATCCTAAAGCCGCAGCTGGAGCGGCGTCGGGGCTACAAGCTCTTCCTGGAC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        644 GTGAACTTCATCCTAAAGCCGCAGCTGGAGCGGCGTCGGGGCTACAAGCTCTTCCTGGAC 703

Qy        601 GACCGCGACCTCCTGCCGCGCGCTGAGCCCTCCGCCGACCTCTTGGTGAACCTGAGCCGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        704 GACCGCGACCTCCTGCCGCGCGCTGAGCCCTCCGCCGACCTCTTGGTGAACCTGAGCCGC 763

Qy        661 TGCCGACGCCTCATCGTGGTGCTTTCGGACGCCTTCCTGAGCCGGGCCTGGTGCAGCCAC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        764 TGCCGACGCCTCATCGTGGTGCTTTCGGACGCCTTCCTGAGCCGGGCCTGGTGCAGCCAC 823

Qy        721 AGCTTCCGGGAGGGCCTGTGCCGGCTGCTGGAGCTCACCCGCAGACCCATCTTCATCACC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        824 AGCTTCCGGGAGGGCCTGTGCCGGCTGCTGGAGCTCACCCGCAGACCCATCTTCATCACC 883

Qy        781 TTCGAGGGCCAGAGGCGCGACCCCGCGCACCCGGCGCTCCGCCTGCTGCGCCAGCACCGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        884 TTCGAGGGCCAGAGGCGCGACCCCGCGCACCCGGCGCTCCGCCTGCTGCGCCAGCACCGC 943

Qy        841 CACCTGGTGACCTTGCTGCTCTGGAGGCCCGGCTCCGTGACTCCTTCCTCCGATTTTTGG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        944 CACCTGGTGACCTTGCTGCTCTGGAGGCCCGGCTCCGTGACTCCTTCCTCCGATTTTTGG 1003

Qy        901 AAAGAAGTGCAGCTGGCGCTGCCGCGGAAGGTGCGGTACAGGCCGGTGGAAGGAGACCCC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1004 AAAGAAGTGCAGCTGGCGCTGCCGCGGAAGGTGCGGTACAGGCCGGTGGAAGGAGACCCC 1063

Qy        961 CAGACGCAGCTGCAGGACGACAAGGACCCCATGCTGATTCTTCGAGGCCGAGTCCCTGAG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1064 CAGACGCAGCTGCAGGACGACAAGGACCCCATGCTGATTCTTCGAGGCCGAGTCCCTGAG 1123

Qy       1021 GGCCGGGCCCTGGACTCAGAGGTGGACCCGGACCCTGAGGGCGACCTGGGTGTCC 1075
              ||||||||||||||||||||||||||||||||||||||||||||||||||| | ||
Db       1124 GGCCGGGCCCTGGACTCAGAGGTGGACCCGGACCCTGAGGGCGACCTGGGTATGC 1178


RESULT 10
BC010806
LOCUS       BC010806                1568 bp    mRNA    linear   ROD 29-JUN-2004
DEFINITION  Mus musculus expressed sequence AI256711, mRNA (cDNA clone
            MGC:18899 IMAGE:4240425), complete cds.
ACCESSION   BC010806
VERSION     BC010806.1  GI:14789848
KEYWORDS    MGC.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1568)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
```

```
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
   PUBMED   12477932
REFERENCE   2  (bases 1 to 1568)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-JUL-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Jeffrey E. Green, M.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Sequencing Group at the Stanford Human Genome
            Center, Stanford University School of Medicine, Stanford, CA  94305
            Web site:       http://www-shgc.stanford.edu
            Contact:  (Dickson, Mark) mcd@paxil.stanford.edu
            Dickson, M., Schmutz, J., Grimwood, J., Rodriquez, A., and Myers,
            R. M.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 25 Row: p Column: 10
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 12746439.
FEATURES             Location/Qualifiers
     source          1. .1568
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="MGC:18899 IMAGE:4240425"
                     /tissue_type="Kidney, normal. 5 month old male mouse."
                     /clone_lib="NCI_CGAP_Kid14"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     gene            1. .1568
                     /gene="AI256711"
                     /note="synonyms: Sigirr, TIR8"
                     /db_xref="LocusID:24058"
                     /db_xref="MGI:1344402"
     CDS             120. .1349
                     /gene="AI256711"
                     /codon_start=1
                     /product="single Ig IL-1 receptor related protein"
                     /protein_id="AAH10806.1"
                     /db_xref="GI:14789849"
                     /db_xref="LocusID:24058"
                     /db_xref="MGI:1344402"
                     /translation="MAGVCDMAPNFLSPSEDQALGLALGREVALNCTAWVFSRPQCPQ
                     PSVQWLKDGLALGNGSHFSLHEDFWVSANPSEIVSSVLVLNLTNAEDYGTFTCSVWNV
                     SSHSFTLWRAGPAGHVAAVLASLLVLVVLLLVALLYVKCRLNMLLWYQDTYGEVEMND
                     GKLYDAYVSYSDCPEDRKFVNFILKPQLERRRGYKLFLEDRDLLPRAEPSADLLVNLS
                     RCRRLIVVLSDAFLSRPWCSQSFREGLCRLLELTRRPIFIFTFEGQRREPIHPALRLLR
                     QHRHLVTLVLWKPGSVTPSSDFWKELQLALPRKVQYRPVEGDPQTRLQDDKDPMLIVR
                     GRAAQGRGMESELDPDPEGDLGVRGPVFGEPPTPLQETRICIGESHGSEMDVSDLGSR
                     NYSARTDFYCLVSEDDV"
ORIGIN

  Query Match             69.8%;  Score 860.2;  DB 10;  Length 1568;
  Best Local Similarity   81.7%;  Pred. No. 2.9e-124;
  Matches 1007;  Conservative    0;  Mismatches  223;  Indels    3;  Gaps    1;

Qy          1 ATGCCAGGTGTCTGTGATAGGGCCCCTGACTTCCTCTCCCCGTCTGAAGACCAGGTGCTG 60
              ||| |||||||||||||| | ||||||| | ||||| ||||| ||||||||||||||   ||
Db        120 ATGGCAGGTGTCTGTGACATGGCCCCTAATTTCCTTTCCCCATCTGAAGACCAGGCCTTG 179

Qy         61 AGGCCTGCCTTGGGCAGCTCAGTGGCTCTGAACTGCACGGCTTGGGTAGTCTCTGGGCCC 120
```

```
                 | | |||| | |||||   ||| ||| |||| ||||| ||||||||  ||||| |||||
Db        180 GGTCTTGCCCTTGGCAGAGAAGTTGCTTTGAATTGCACAGCTTGGGTGTTCTCTAGGCCC 239

Qy        121 CACTGCTCCCTGCCTTCAGTCCAGTGGCTGAAAGACGGGCTTCCATTGGGAATTGGGGGC 180
              || ||  ||| ||| ||||| ||||||||||||||| || ||  |||||||  | |||  ||
Db        240 CAGTGTCCCCAGCCATCAGTGCAGTGGCTGAAAGATGGTCTGGCATTGGGCAATGGAAGC 299

Qy        181 CACTACAGCCTCCACGAGTACTCCTGGGTCAAGGCCAACCTGTCAGAGGTGCTTGTGTCC 240
              |||| ||||||||| ||| ||| |||||||  ||||||  | |||||   | ||||||||
Db        300 CACTTCAGCCTCCATGAGGACTTCTGGGTCAGCGCCAACTTCTCAGA---GATTGTGTCC 356

Qy        241 AGTGTCCTGGGGGTCAACGTGACCAGCACTGAAGTCTATGGGGCCTTCACCTGCTCCATC 300
              |||||||||| | ||||| ||||||    | || | ||||||  |||||||||| ||  ||
Db        357 AGTGTCCTGGTGCTCAACTTGACCAATGCAGAGGACTATGGAACCTTCACCTGTTCTGTC 416

Qy        301 CAGAACATCAGCTTCTCCTCCTTCACTCTTCAGAGAGCTGGCCCTACAAGCCACGTGGCT 360
                |||  |||||| |   |||||||||||  | ||||||||||| |  |||| ||||||
Db        417 TGGAATGTCAGCTCCCATTCCTTCACTCTTTGGCGAGCTGGCCCTGCTGGCCATGTGGCT 476

Qy        361 GCGGTGCTGGCCTCCCTCCTGGTCCTGCTGGCCCTGCTGCTGGCCGCCCTGCTCTATGTC 420
              || || ||||| |||||||||||||||| |||  ||||||||||  ||||||||||||||
Db        477 GCAGTACTGGCTTCCCTCCTGGTCCTGGTGGTTCTGCTGCTGGTGGCCCTGCTCTATGTT 536

Qy        421 AAGTGCCGTCTCAACGTGCTGCTCTGGTACCAGGACGCGTATGGGGAGGTGGAGATAAAC 480
              ||||| || || ||| ||||||| |||||||| ||| | || ||||||||||||||| |||
Db        537 AAGTGTCGGCTGAACATGCTGCTTTGGTACCAAGACACTTACGGGGAGGTGGAGATGAAC 596

Qy        481 GACGGGAAGCTCTACGACGCCTACGTCTCCTACAGCGACTGCCCCGAGGACCGCAAGTTC 540
              || |||||| | ||||| |||||||| ||||| |||||||||| ||||||||||| ||
Db        597 GATGGGAAGTTATACGATGCCTACGTGTCCTATAGCGACTGCCCAGAGGACCGCAAATTT 656

Qy        541 GTGAACTTCATCCTAAAGCCGCAGCTGGAGCGGCGTCGGGGCTACAAGCTCTTCCTGGAC 600
              || || || || || ||||| ||| |||||||||||||||| ||||| |||||||| ||
Db        657 GTAAATTTTATTCTGAAGCCTCAGTTGGAGCGGCGTCGGGGATACAAACTCTTCCTAGAG 716

Qy        601 GACCGCGACCTCCTGCCGCGCGCTGAGCCCTCCGCCGACCTCTTGGTGAACCTGAGCCGC 660
              |||||||||||| |||| ||||| |||||||| |||||||| |||||||||||||| |||
Db        717 GACCGCGACCTCTTGCCTCGCGCGGAGCCCTCTGCCGACCTTTTGGTGAACCTGAGTCGC 776

Qy        661 TGCCGACGCCTCATCGTGGTGCTTTCGGACGCCTTCCTGAGCCGGGCCTGGTGCAGCCAC 720
              || || || |||||||||||| ||||| || |||||||| |||||| |||||| |||||
Db        777 TGTCGGCGTCTCATCGTGGTTCTTTCAGATGCCTTCCTAAGCCGGCCCTGGTGTAGCCAG 836

Qy        721 AGCTTCCGGGAGGGCCTGTGCCGGCTGCTGGAGCTCACCCGCAGACCCATCTTCATCACC 780
              |||||||||||||| |||||||| || |||||||||||||||||||| ||||||||||||
Db        837 AGCTTCCGGGAGGGACTGTGCCGCCTACTGGAGCTCACCCGCAGACCTATCTTCATCACC 896

Qy        781 TTCGAGGGCCAGAGGCGCGACCCCGCGCACCCGGCGCTCCGCCTGCTGCGCCAGCACCGC 840
              || |||||||||||||| || |||   ||||| || ||||| || ||||||||||||||||
Db        897 TTTGAGGGCCAGAGGCGTGAGCCCATACACCCTGCTCTCCGGCTCCTGCGCCAGCACCGC 956

Qy        841 CACCTGGTGACCTTGCTGCTCTGGAGGCCCGGCTCCGTGACTCCTTCCTCCGATTTTTGG 900
              ||||| |||||| || |||| |||| ||| ||||||||||||||||||| |||||||||
Db        957 CACCTCGTGACCCTGGTGCTTTGGAAGCCTGGCTCCGTGACTCCTTCCTCTGATTTTTGG 1016

Qy        901 AAAGAAGTGCAGCTGGCGCTGCCGCGGAAGGTGCGGTACAGGCCGGTGGAAGGAGACCCC 960
              |||||  | ||||| || ||||| |||||||||| |||||||||||||| ||||||||
Db       1017 AAAGAGCTACAGCTAGCACTGCCACGGAAGGTGCAGTACAGGCCGGTGGAGGGAGACCCT 1076

Qy        961 CAGACGCAGCTGCAGGACGACAAGGACCCCATGCTGATTCTTCGAGGCCGAGTCCCTGAG 1020
              || || |  || ||||| ||||| || |||||||| ||  |  |||| || |   |  ||
Db       1077 CAAACCCGACTTCAGGATGACAAAGATCCCATGCTAATCGTGAGAGGACGTGCTGCCCAG 1136

Qy       1021 GGCCGGGCCCTGGACTCAGAGGTGGACCCGGACCCTGAGGGCGACCTGGGTGTCCGGGGG 1080
              ||||||| | |||| |||||| |||| || |||||||||| |||||||||||||| ||
Db       1137 GGCCGGGGCATGGAGTCAGAGCTGGATCCAGACCCTGAGGGAGACCTGGGTGTCCGTGGA 1196

Qy       1081 CCTGTTTTTGGAGAGCCATCAGCTCCACCGCACACCAGTGGGGTCTCGCTGGGAGAGAGC 1140
              ||||| ||||| |||||| || |||||| |||     |  || |||   | |||||||||
Db       1197 CCTGTCTTTGGGGAGCCACCAACTCCACTGCAGGAAACCAGGATCTGCATAGGAGAGAGC 1256

Qy       1141 CGGAGCAGCGAAGTGGACGTCTCGGATCTCGGCTCGCGAAACTACAGTGCCCGCACAGAC 1200
              |   |||| ||| |||| ||||| || |||||||| |||||||||||||| || ||||||
```

```
Db       1257 CACGGCAGTGAAATGGATGTCTCTGACCTCGGCTCTCGAAACTACAGTGCACGGACAGAC 1316

Qy       1201 TTCTACTGCCTGGTGTCCAAGGATGATATGTAG 1233
             ||||||||||| |||||  |||||||| |||||
Db       1317 TTCTACTGCCTCGTGTCTGAGGATGATGTGTAG 1349


RESULT 11
AF239957
LOCUS       AF239957               1421 bp    mRNA    linear   ROD 21-MAR-2000
DEFINITION  Mus musculus single Ig IL-1R-related protein mRNA, complete cds.
ACCESSION   AF239957
VERSION     AF239957.1  GI:7271942
KEYWORDS    .
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1421)
  AUTHORS   Poltorak,A. and Beutler,B.
  TITLE     Molecular cloning of mouse SIGIRR, an orphan receptor with a
            Toll-like cytoplasmic domain
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1421)
  AUTHORS   Poltorak,A. and Beutler,B.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-MAR-2000) Internal Medicine, UT Southwestern Medical
            Center/HHMI, 5323 Harry Hines Boulevard, Dallas, TX 75235-9050, USA
FEATURES             Location/Qualifiers
     source          1. .1421
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="BALB/c"
                     /db_xref="taxon:10090"
                     /cell_line="RAW 264.7"
     CDS             109. .1338
                     /note="SIGIRR; Toll-like cytoplasmic domain"
                     /codon_start=1
                     /product="single Ig IL-1R-related protein"
                     /protein_id="AAF44701.1"
                     /db_xref="GI:7271943"
                     /translation="MAGVCDMAPNFLSPSEDQALGLALGREVALNCTAWVFSRPQCPQ
                     PSVQWLKDGLALGNGSHFSLHEDFWVSANFSEIVSSVLVLNLTNAEDYGTFTCSVWNV
                     SSHSFTLWRAGPAGHVAAVLASLLVLVVLLLVALLYVKCRLNMLLWYQDTYGEVEMND
                     GKLYDAYVSYSDCPEDRKFVNFILKPQLERRRGYKPFLEDRDLLPRAEPSADLLVNLS
                     RCRRLIVVLSDAFLSRPWCSQSFREGLCRLLELTRRPIFITFEGQRREPIHPALRLLR
                     QHRHLVTLVLWKPGSVTPSSDFWKELQLALPRKVQYRPVEGDPQTRLQDDKDPMLIVR
                     GRAAQGRGMESELDPDPEGDLGVRGPVFGEPPTPLQETRICIGESHGSEMDVSDLGSR
                     NYSARTDFYCLVSEDDV"

ORIGIN

  Query Match           69.5%;  Score 856.6;  DB 10;  Length 1421;
  Best Local Similarity  81.5%;  Pred. No. 1.1e-123;
  Matches 1004;  Conservative    0;  Mismatches  225;  Indels    3;  Gaps    1;

Qy          1 ATGCCAGGTGTCTGTGATAGGGCCCCTGACTTCCTCTCCCCGTCTGAAGACCAGGTGCTG 60
              ||| |||||||||||| | ||||||| | ||||| ||||| ||||||||||||    ||
Db        109 ATGGCAGGTGTCTGTGACATGGCCCCTAATTTCCTTTCCCCATCTGAAGACCAGGCCTTG 168

Qy         61 AGGCCTGCCTTGGGCAGCTCAGTGGCTCTGAACTGCACGGCTTGGGTAGTCTCTGGGCCC 120
               | | |||| | |||||   ||| ||| |||| ||||| ||||||||  ||||| |||||
Db        169 GGTCTTGCCCTTGGCAGAGAAGTTGCTTTGAATTGCACAGCTTGGGTGTTCTCTAGGCCC 228

Qy        121 CACTGCTCCCTGCCTTCAGTCCAGTGGCTGAAAGACGGGCTTCCATTGGGAATTGGGGGC 180
              || ||  ||| ||| ||||| ||||||||||||||| || ||  ||||||| | |||  ||
Db        229 CAGTGTCCCCAGCCATCAGTGCAGTGGCTGAAAGATGGTCTGGCATTGGGCAATGGAAGC 288

Qy        181 CACTACAGCCTCCACGAGTACTCCTGGGTCAAGGCCAACCTGTCAGAGGTGCTTGTGTCC 240
              |||| ||||||||| ||| ||| ||||||||  |||||| | |||||    | ||||||||
Db        289 CACTTCAGCCTCCATGAGGACTTCTGGGTCAGCGCCAACTTCTCAGA---GATTGTGTCC 345

Qy        241 AGTGTCCTGGGGGGTCAACGTGACCAGCACTGAAGTCTATGGGGCCTTCACCTGCTCCATC 300
              |||||||||| | ||||| ||||||   | || | ||||||  |||||||||| ||  ||

Db        346 AGTGTCCTGGTGCTCAACTTGACCAATGCAGAGGACTATGGAACCTTCACCTGTTCTGTC 405

Qy        301 CAGAACATCAGCTTCTCCTCCTTCACTCTTCAGAGAGCTGGCCCTACAAGCCACGTGGCT 360
              |||  |||||| |   |||||||||||  | ||||||||||| |  |||| ||||||
Db        406 TGGAATGTCAGCTCCCATTCCTTCACTCTTTGGCGAGCTGGCCCTGCTGGCCATGTGGCT 465

Qy        361 GCGGTGCTGGCCTCCCTCCTGGTCCTGCTGGCCCTGCTGCTGGCCGCCCTGCTCTATGTC 420
              || || ||||| |||||||||||||| |||  |||||||||  |||||||||||||||
Db        466 GCAGTACTGGCTTCCCTCCTGGTCCTGGTGGTTCTGCTGCTGGTGGCCCTGCTCTATGTT 525

Qy        421 AAGTGCCGTCTCAACGTGCTGCTCTGGTACCAGGACGCGTATGGGGAGGTGGAGATAAAC 480
              ||||| || || ||| ||||||| |||||||| ||| | || ||||||||||||||| |||
Db        526 AAGTGTCGGCTGAACATGCTGCTTTGGTACCAAGACACTTACGGGGAGGTGGAGATGAAC 585

Qy        481 GACGGGAAGCTCTACGACGCCTACGTCTCCTACAGCGACTGCCCCGAGGACCGCAAGTTC 540
              || |||||| | ||||| |||||||| ||||| |||||||||| |||||||||| ||
Db        586 GATGGGAAGTTATACGATGCCTACGTGTCCTATAGCGACTGCCCAGAGGACCGCAAATTT 645

Qy        541 GTGAACTTCATCCTAAAGCCGCAGCTGGAGCGGCGTCGGGGCTACAAGCTCTTCCTGGAC 600
              || || || || || ||||| ||| |||||||||||||||| ||||| | |||||| ||
Db        646 GTAAATTTTATTCTGAAGCCTCAGTTGGAGCGGCGTCGGGGATACAAACCCTTCCTAGAG 705

Qy        601 GACCGCGACCTCCTGCCGCGCGCTGAGCCCTCCGCCGACCTCTTGGTGAACCTGAGCCGC 660
              |||||||||||| |||| ||||| |||||||| |||||||| |||||||||||||| |||
Db        706 GACCGCGACCTCTTGCCTCGCGCGGAGCCCTCTGCCGACCTTTTGGTGAACCTGAGTCGC 765

Qy        661 TGCCGACGCCTCATCGTGGTGCTTTCGGACGCCTTCCTGAGCCGGGCCTGGTGCAGCCAC 720
              || || || |||||||||| ||||| || |||||||| |||||| ||||||| |||||
Db        766 TGTCGGCGTCTCATCGTGGTTCTTTCAGATGCCTTCCTAAGCCGGCCCTGGTGTAGCCAG 825

Qy        721 AGCTTCCGGGAGGGCCTGTGCCGGCTGCTGGAGCTCACCCGCAGACCCATCTTCATCACC 780
              |||||||||||||| |||||||| || |||||||||||||||||||| ||||||||||||
Db        826 AGCTTCCGGGAGGGACTGTGCCGCCTACTGGAGCTCACCCGCAGACCTATCTTCATCACC 885

Qy        781 TTCGAGGGCCAGAGGCGCGACCCCGCGCACCCGGCGCTCCGCCTGCTGCGCCAGCACCGC 840
              || |||||||||||||| || |||    ||||| || ||||| || ||||||||||||||
Db        886 TTTGAGGGCCAGAGGCGTGAGCCCATACACCCTGCTCTCCGGCTCCTGCGCCAGCACCGC 945

Qy        841 CACCTGGTGACCTTGCTGCTCTGGAGGCCCGGCTCCGTGACTCCTTCCTCCGATTTTTGG 900
              ||||| |||||| || |||| |||| ||| ||||||||||||||||||| |||||||||
Db        946 CACCTCGTGACCCTGGTGCTTTGGAAGCCTGGCTCCGTGACTCCTTCCTCTGATTTTTGG 1005

Qy        901 AAAGAAGTGCAGCTGGCGCTGCCGCGGAAGGTGCGGTACAGGCCGGTGGAAGGAGACCCC 960
              |||||  | ||||| || ||||| |||||||||| |||||||||||||| ||||||||
Db       1006 AAAGAGCTACAGCTAGCACTGCCACGGAAGGTGCAGTACAGGCCGGTGGAGGGAGACCCT 1065

Qy        961 CAGACGCAGCTGCAGGACGACAAGGACCCCATGCTGATTCTTCGAGGCCGAGTCCCTGAG 1020
              || || |  || ||||| ||||| || |||||||| ||  |  |||| || |   | ||
Db       1066 CAAACCCGACTTCAGGATGACAAAGATCCCATGCTAATCGTGAGAGGACGTGCTGCCCAG 1125

Qy       1021 GGCCGGGCCCTGGACTCAGAGGTGGACCCGGACCCTGAGGGCGACCTGGGTGTCCGGGGG 1080
              ||||||| | |||| |||||| |||| || |||||||||| |||||||||||||| ||
Db       1126 GGCCGGGGCATGGAGTCAGAGCTGGATCCAGACCCTGAGGGAGACCTGGGTGTCCGTGGA 1185

Qy       1081 CCTGTTTTTGGAGAGCCATCAGCTCCACCGCACACCAGTGGGGTCTCGCTGGGAGAGAGC 1140
              ||||| ||||| |||||| || |||||| |||     |   || |||   | |||||||||
Db       1186 CCTGTCTTTGGGGAGCCACCAACTCCACTGCAGGAAACCAGGATCTGCATAGGAGAGAGC 1245

Qy       1141 CGGAGCAGCGAAGTGGACGTCTCGGATCTCGGCTCGCGAAACTACAGTGCCCGCACAGAC 1200
              |   |||| ||| |||| ||||| || |||||||| || |||||||||| || ||||||
Db       1246 CACGGCAGTGAAATGGATGTCTCTGACCTCGGCTCTCGNAACTACAGTGCACGGACAGAC 1305

Qy       1201 TTCTACTGCCTGGTGTCCAAGGATGATATGTA 1232
              |||||||||| |||||  |||||||| ||||
Db       1306 TTCTACTGCCTCGTGTCTGAGGATGATGTGTA 1337


RESULT 12
AF113795
LOCUS       AF113795               1540 bp    mRNA    linear   ROD 10-MAR-2004
DEFINITION  Mus musculus toll/interleukin-1 receptor 8 (TIR8) mRNA, complete
            cds.
ACCESSION   AF113795
```

```
VERSION     AF113795.1  GI:6714887
KEYWORDS    .
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1540)
  AUTHORS   Garlanda,C., Riva,F., Polentarutti,N., Buracchi,C., Sironi,M., De
            Bortoli,M., Muzio,M., Bergottini,R., Scanziani,E., Vecchi,A.,
            Hirsch,E. and Mantovani,A.
  TITLE     Intestinal inflammation in mice deficient in Tir8, an inhibitory
            member of the IL-1 receptor family
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 101 (10), 3522-3526 (2004)
   PUBMED   14993616
REFERENCE   2  (bases 1 to 1540)
  AUTHORS   Penton-Rol,G., Polentarutti,N., Muzio,M. and Mantovani,A.
  TITLE     TIR8: a novel toll/interleukin-1 receptor family member
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 1540)
  AUTHORS   Penton-Rol,G., Polentarutti,N., Muzio,M. and Mantovani,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-DEC-1998) Immunology, Mario Negri Institute, Via
            Eritrea 62, Milano, MI I-20152, Italy
FEATURES             Location/Qualifiers
     source          1. .1540
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
     gene            1. .1540
                     /gene="TIR8"
     CDS             115. .1344
                     /gene="TIR8"
                     /note="putative transmembrane receptor of the Toll/IL-1
                     family; cytoplasmic domain similar to other IL-1R family
                     members"
                     /codon_start=1
                     /product="toll/interleukin-1 receptor 8"
                     /protein_id="AAF26200.1"
                     /db_xref="GI:6714888"
                     /translation="MAGVCDMAPNFLSPSEDQALGLALGREVALNCTAWVFSRPQCPQ
                     PSVQWLKDGLALGNGSHFSLHEDFWVSANFSEIVSSVLVLNLTNAEDYGTFTCSVWNV
                     SSHSFTLWRAGPAGHVAAVLASLLVLVVLLLVALLYVKCRLNMLLWYQDTYREVEMND
                     GKLYDAYVSYSDCPEDRKFVNFILKPQLERRRGYKLFLEDRDLLPRADASADLLVNLS
                     RCRRLIVVLSDAFLSRPWCSQSFREDCAAYWSSPADLSSSPLRARGVSPYTLLSGSCR
                     HDRHLVTLVLWKPGSVTPSSDFWKELQLALPRKVQYRPVEGDPQTRLQDDKDPMLIVR
                     GRAAQGRGMESELDPDPEGDLGVRGPVFGEPPTPLQETRICIGESHGSEMDVSDLGSR
                     NYSARTDFYCLVSEDDV"
ORIGIN

  Query Match           67.0%;  Score 826;  DB 10;  Length 1540;
  Best Local Similarity  81.0%;  Pred. No. 6.1e-119;
  Matches  999;  Conservative    0;  Mismatches  230;  Indels    5;  Gaps    3;

Qy          1 ATGCCAGGTGTCTGTGATAGGGCCCCTGACTTCCTCTCCCCGTCTGAAGACCAGGTGCTG 60
              ||| |||||||||||||| | ||||||| | ||||| ||||| ||||||||||||||   ||
Db        115 ATGGCAGGTGTCTGTGACATGGCCCCTAATTTCCTTTCCCCATCTGAAGACCAGGCCTTG 174

Qy         61 AGGCCTGCCTTGGGCAGCTCAGTGGCTCTGAACTGCACGGCTTGGGTAGTCTCTGGGCCC 120
                | | |||| | |||||   ||| ||| |||| ||||| ||||||||  ||||| |||||
Db        175 GGTCTTGCCCTTGGCAGAGAAGTTGCTTTGAATTGCACAGCTTGGGTGTTCTCTAGGCCC 234

Qy        121 CACTGCTCCCTGCCTTCAGTCCAGTGGCTGAAAGACGGGCTTCCATTGGGAATTGGGGGC 180
              || ||  ||| ||| ||||| ||||||||||||||| || ||  ||||||| | |||  ||
Db        235 CAGTGTCCCCAGCCATCAGTGCAGTGGCTGAAAGATGGTCTGGCATTGGGCAATGGAAGC 294

Qy        181 CACTACAGCCTCCACGAGTACTCCTGGGTCAAGGCCAACCTGTCAGAGGTGCTTGTGTCC 240
              |||| ||||||||| ||| ||| ||||||||  |||||| | |||||    | ||||||||
Db        295 CACTTCAGCCTCCATGAGGACTTCTGGGTCAGCGCCAACTTCTCAGA---GATTGTGTCC 351

Qy        241 AGTGTCCTGGGGGTCAACGTGACCAGCACTGAAGTCTATGGGGCCTTCACCTGCTCCATC 300
              |||||||||| | ||||| ||||||    | || | ||||||  |||||||||| ||  ||
Db        352 AGTGTCCTGGTGCTCAACTTGACCAATGCAGAGGACTATGGAACCTTCACCTGTTCTGTC 411

Qy        301 CAGAACATCAGCTTCTCCTCCTTCACTCTTCAGAGAGCTGGCCCTACAAGCCACGTGGCT 360
              |||  |||||| |   |||||||||||  | |||| |||||| |  | || ||||||
Db        412 TGGAATGTCAGCTCCCATTCCTTCACTCTTTGGCGAGCCGGCCCTGCTGGGCATGTGGCT 471

Qy        361 GCGGTGCTGGCCTCCCTCCTGGTCCTGCTGGCCCTGCTGCTGGCCGCCCTGCTCTATGTC 420
              || || ||||| |||||||||||||| |||  |||||||||  |||||||||||||
Db        472 GCAGTACTGGCTTCCCTCCTGGTCCTGGTGGTTCTGCTGCTGGTGGCCCTGCTCTATGTT 531

Qy        421 AAGTGCCGTCTCAACGTGCTGCTCTGGTACCAGGACGCGTATGGGGAGGTGGAGATAAAC 480
              ||||| || || ||| ||||||| |||||||| ||| | ||  |||||||||||| |||
Db        532 AAGTGTCGGCTGAACATGCTGCTTTGGTACCAAGACACTTACCGGGAGGTGGAGATGAAC 591

Qy        481 GACGGGAAGCTCTACGACGCCTACGTCTCCTACAGCGACTGCCCCGAGGACCGCAAGTTC 540
              || |||||| | ||||| |||||||| ||||| |||||||||| |||||||||| ||
Db        592 GATGGGAAGTTATACGATGCCTACGTGTCCTATAGCGACTGCCCAGAGGACCGCAAATTT 651

Qy        541 GTGAACTTCATCCTAAAGCCGCAGCTGGAGCGGCGTCGGGGCTACAAGCTCTTCCTGGAC 600
              || || || || || ||||| ||| |||||||||||||||| ||||| |||||||| ||
Db        652 GTAAATTTTATTCTGAAGCCTCAGTTGGAGCGGCGTCGGGGATACAAACTCTTCCTAGAG 711

Qy        601 GACCGCGACCTCCTGCCGCGCGCTGAGCCCTCCGCCGACCTCTTGGTGAACCTGAGCCGC 660
              |||||||||||| |||| ||||| ||  |||| |||||||| |||||||||||||| |||
Db        712 GACCGCGACCTCTTGCCTCGCGCGGACGCCTCTGCCGACCTTTTGGTGAACCTGAGTCGC 771

Qy        661 TGCCGACGCCTCATCGTGGTGCTTTCGGACGCCTTCCTGAGCCGGGCCTGGTGCAGCCAC 720
              || || || |||||||||| |||| || |||||||| |||||| |||||| |||||
Db        772 TGTCGGCGTCTCATCGTGGTTCTTTCAGATGCCTTCCTAAGCCGGCCCTGGTGTAGCCAG 831

Qy        721 AGCTTCCGGGAGGGCCTGTGCCGGCTGCTGGAGCTCACCCGCAGACCCATCTTCATCACC 780
              |||||||||||| || |||||||| || ||||||||||||||||||| ||||||||||||
Db        832 AGCTTCCGGGA-GGACTGTGCCGCCTACTGGAGCTCACCCGCAGACCTATCTTCATCACC 890

Qy        781 TTCGAGGGCCAGAGGCGCGACCCCGCGCACCCGGCGCTCCGCCTGCTG-CGCCAGCACCG 839
              || |||||||||||||| || |||   ||||| || ||||| || ||| |||||  ||||
Db        891 TTTGAGGGCCAGAGGCGTGAGCCCATACACCCTGCTCTCCGGCTCCTGCCGCCACGACCG 950

Qy        840 CCACCTGGTGACCTTGCTGCTCTGGAGGCCCGGCTCCGTGACTCCTTCCTCCGATTTTTG 899
              |||||| |||||| || |||| |||| ||| ||||||||||||||||||| ||||||||
Db        951 CCACCTCGTGACCCTGGTGCTTTGGAAGCCTGGCTCCGTGACTCCTTCCTCTGATTTTTG 1010

Qy        900 GAAAGAAGTGCAGCTGGCGCTGCCGCGGAAGGTGCGGTACAGGCCGGTGGAAGGAGACCC 959
              ||||||  | ||||| || |||||||||||| ||||||||||||||| ||||||||
Db       1011 GAAAGAGCTACAGCTAGCACTGCCACGGAAGGTGCAGTACAGGCCGGTGGAGGGAGACCC 1070

Qy        960 CCAGACGCAGCTGCAGGACGACAAGGACCCCATGCTGATTCTTCGAGGCCGAGTCCCTGA 1019
              || || |  || ||||| ||||| || |||||||| ||  |  |||| || |  |  |
Db       1071 TCAAACCCGACTTCAGGATGACAAAGATCCCATGCTAATCGTGAGAGGACGTGCTGCCCA 1130

Qy       1020 GGGCCGGGCCCTGGACTCAGAGGTGGACCCGGACCCTGAGGGCGACCTGGGTGTCCGGGG 1079
              |||||||| | |||| |||||| |||| || |||||||||| |||||||||||||| ||
Db       1131 GGGCCGGGGCATGGAGTCAGAGCTGGATCCAGACCCTGAGGGAGACCTGGGTGTCCGTGG 1190

Qy       1080 GCCTGTTTTTGGAGAGCCATCAGCTCCACCGCACACCAGTGGGGTCTCGCTGGGAGAGAG 1139
              ||||| ||||| |||||| || |||||| |||     ||  |||   | |||||||||
Db       1191 ACCTGTCTTTGGGGAGCCACCAACTCCACTGCAGGAAACCAGGATCTGCATAGGAGAGAG 1250

Qy       1140 CCGGAGCAGCGAAGTGGACGTCTCGGATCTCGGCTCGCGAAACTACAGTGCCCGCACAGA 1199
              ||   |||| ||| |||| ||||| || |||||||| |||||||||||||| || |||||
Db       1251 CCACGGCAGTGAAATGGATGTCTCTGACCTCGGCTCTCGAAACTACAGTGCACGGACAGA 1310

Qy       1200 CTTCTACTGCCTGGTGTCCAAGGATGATATGTAG 1233
              |||||||||||| |||||  |||||||| |||||
Db       1311 CTTCTACTGCCTCGTGTCTGAGGATGATGTGTAG 1344


RESULT 13
AK093427
LOCUS       AK093427               1867 bp    mRNA    linear   PRI 30-JAN-2004
DEFINITION  Homo sapiens cDNA FLJ36108 fis, clone TESTI2021895, moderately
            similar to Mus musculus toll/interleukin-1 receptor 8 (TIR8) mRNA.
ACCESSION   AK093427
VERSION     AK093427.1  GI:21752294
KEYWORDS    oligo capping; fis (full insert sequence).
```

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Ota,T., Suzuki,Y., Nishikawa,T., Otsuki,T., Sugiyama,T., Irie,R., Wakamatsu,A., Hayashi,K., Sato,H., Nagai,K., Kimura,K., Makita,H., Sekine,M., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T., Ishii,S., Yamamoto,J., Saito,K., Kawai,Y., Isono,Y., Nakamura,Y., Nagahari,K., Murakami,K., Yasuda,T., Iwayanagi,T., Wagatsuma,M., Shiratori,A., Sudo,H., Hosoiri,T., Kaku,Y., Kodaira,H., Kondo,H., Sugawara,M., Takahashi,M., Kanda,K., Yokoi,T., Furuya,T., Kikkawa,E., Omura,Y., Abe,K., Kamihara,K., Katsuta,N., Sato,K., Tanikawa,M., Yamazaki,M., Ninomiya,K., Ishibashi,T., Yamashita,H., Murakawa,K., Fujimori,K., Tanai,H., Kimata,M., Watanabe,M., Hiraoka,S., Chiba,Y., Ishida,S., Ono,Y., Takiguchi,S., Watanabe,S., Yosida,M., Hotuta,T., Kusano,J., Kanehori,K., Takahashi-Fujii,A., Hara,H., Tanase,T., Nomura,Y., Togiya,S., Komai,F., Hara,R., Takeuchi,K., Arita,M., Imose,N., Musashino,K., Yuuki,H., Oshima,A., Sasaki,N., Aotsuka,S., Yoshikawa,Y., Matsunawa,H., Ichihara,T., Shiohata,N., Sano,S., Moriya,S., Momiyama,H., Satoh,N., Takami,S., Terashima,Y., Suzuki,O., Nakagawa,S., Senoh,A., Mizoguchi,H., Goto,Y., Shimizu,F., Wakebe,H., Hishigaki,H., Watanabe,T., Sugiyama,A., Takemoto,M., Kawakami,B., Yamazaki,M., Watanabe,K., Kumagai,A., Itakura,S., Fukuzumi,Y., Fujimori,Y., Komiyama,M., Tashiro,H., Tanigami,A., Fujiwara,T., Ono,T., Yamada,K., Fujii,Y., Ozaki,K., Hirao,M., Ohmori,Y., Kawabata,A., Hikiji,T., Kobatake,N., Inagaki,H., Ikema,Y., Okamoto,S., Okitani,R., Kawakami,T., Noguchi,S., Itoh,T., Shigeta,K., Senba,T., Matsumura,K., Nakajima,Y., Mizuno,T., Morinaga,M., Sasaki,M., Togashi,T., Oyama,M., Hata,H., Watanabe,M., Komatsu,T., Mizushima-Sugano,J., Satoh,T., Shirai,Y., Takahashi,Y., Nakagawa,K., Okumura,K., Nagase,T., Nomura,N., Kikuchi,H., Masuho,Y., Yamashita,R., Nakai,K., Yada,T., Nakamura,Y., Ohara,O., Isogai,T. and Sugano,S.
TITLE Complete sequencing and characterization of 21,243 full-length human cDNAs
JOURNAL Nat. Genet. 36 (1), 40-45 (2004)
PUBMED 14702039
REFERENCE 2
AUTHORS Nishi,T., Ota,T., Nakagawa,S., Senoh,A., Mizuguchi,H., Inagaki,H., Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K., Wagatsuma,M., Murakawa,K., Kanehori,K., Takahashi-Fujii,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S., Nagahari,K., Masuho,Y., Nagai,K. and Isogai,T.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 1867)
AUTHORS Isogai,T. and Yamamoto,J.
TITLE Direct Submission
JOURNAL Submitted (04-JUL-2002) Takao Isogai, FLJ Project(HRI Team); 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.
FEATURES Location/Qualifiers
source 1. .1867
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="TESTI2021895"
/tissue_type="testis"
/clone_lib="TESTI2"
/note="cloning vector: pME18SFL3"
ORIGIN

Query Match 45.1%; Score 556.2; DB 9; Length 1867;
Best Local Similarity 70.8%; Pred. No. 5.2e-77;
Matches 994; Conservative 0; Mismatches 3; Indels 406; Gaps 3;

```
Qy      203 CCTGGGTCAAGGCCAACCTGTCAGAGGTGCTTGTGTCCAGTGTCCTGGGGGTCAACGTGA 262
            || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      319 CCAGGGTCAAGGCCAACCTGTCAGAGGTGCTTGTGTCCAGTGTCCTGGGGGTCAACGTGA 378

Qy      263 CCAGCACTGAAGTCTATGGGGCCTTCACCTGCTCCATCCAGAACATCAGCTTCTCCTCCT 322
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      379 CCAGCACTGAAGTCTATGGGGCCTTCACCTGCTCCATCCAGAACATCAGCTTCTCCTCCT 438

Qy      323 TCACTCTTCAGAGAGCTGGCCCTACAAGCCACGTGGCTGCGGTGCTGGCCTCCCTCCTGG 382
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      439 TCACTCTTCAGAGAGCTGGCCCTACAAGCCACGTGGCTGCGGTGCTGGCCTCCCTCCTGG 498

Qy      383 TCCTGCTGGCCCTGCTGCTGGCCGCCCTGCTCTATGTCAAGTGCCGTCTCAACGTGCTGC 442
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      499 TCCTGCTGGCCCTGCTGCTGGCCGCCCTGCTCTATGTCAAGTGCCGTCTCAACGTGCTGC 558

Qy      443 TCTGGTACCAGGACGCGTATGGGGAGGTGGAGATAAACGACGGGAAGCTCTACGACGCCT 502
            ||||||||||||||||||||||||||||||||||||
Db      559 TCTGGTACCAGGACGCGTATGGGGAGGTGGAGATAAAC---------------------- 596

Qy      503 ACGTCTCCTACAGCGACTGCCCCGAGGACCGCAAGTTCGTGAACTTCATCCTAAAGCCGC 562
                    ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      597 ------------GCGACTGCCCCGAGGACCGCAAGTTCGTGAACTTCATCCTAAAGCCGC 644

Qy      563 AGCTGGAGCGGCGTCGGGGCTACAAGCTCTTCCTGGACGACCGCGACCTCCTGCCGCGCG 622
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      645 AGCTGGAGCGGCGTCGGGGCTACAAGCTCTTCCTGGACGACCGCGACCTCCTGCCGCGCG 704

Qy      623 CTGAGCCCTCCGCCGACCTCTTGGTGAACCTGAGCCGCTGCCGACGCCTCATCGTGGTGC 682
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      705 CTGAGCCCTCCGCCGACCTCTTGGTGAACCTGAGCCGCTGCCGACGCCTCATCGTGGTGC 764

Qy      683 TTTCGGACGCCTTCCTGAGCCGGGCCTGGTGCAGCCACAGCTTCC-------------- 727
            ||||||||||||||||||||||||||||||||||||||||||||
Db      765 TTTCGGACGCCTTCCTGAGCCGGGCCTGGTGCAGCCACAGCTTCCGGTGGGTCCCGCGCG 824

Qy      728 ------------------------------------------------------GGGAGGG 734
                                                                   ||| |||
Db      825 GGGTTGGGTGGGCCCCAGCGTAGCACCCACCCCCCTGACGGTCCCGCCCCGCAGGGGGGG 884

Qy      735 CCTGTGCCGGCTGCTGGAGCTCACCCGCAGACCCATCTTCATCACCTTCGAGGGCCAGAG 794
            |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db      885 CCTGTGCCGGCTGCTGGTGCTCACCCGCAGACCCATCTTCATCACCTTCGAGGGCCAGAG 944

Qy      795 GCGCGACCCCGCGCACCCGGCGCTCCGCCTGCTGCGCCAGCACCGCCACCTGGTGACCTT 854
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      945 GCGCGACCCCGCGCACCCGGCGCTCCGCCTGCTGCGCCAGCACCGCCACCTGGTGACCTT 1004

Qy      855 GCTGCTCTGGAGGCCCGGCTCCGT---------------------------------- 878
            |||||||||||||||||||||||
Db      1005 GCTGCTCTGGAGGCCCGGCTCCGTGGTGCGGAGCAGGCGCGGGAGGGTCCGGGGCTAGCG 1064

Qy      879 ---------------------------------------------------------- 878

Db      1065 GCGGGTTAGAGATGGGCGGTGCCCGGGCTCCAGGCTGGGACCCCTCCGTGGGGAGCTCTG 1124

Qy      879 ---------------------------------------------------------- 878

Db      1125 CGGCACCACGCTTTGTGAATGGGCCCTGGGGGGAGGTTCCGCTGCCTGGGGCCCCGATGC 1184

Qy      879 ---------------------------------------------------------- 878

Db      1185 GGGGAGCCGCCCTTGAGGCCCCCGGAGCCACGGAATAGCTGTCGCAGGGCGTGGAACCCG 1244

Qy      879 ---------------------------------------------------------- 878

Db      1245 TGGGCAGCCGCAGGTGTGCTCTTGGGGGCCAGGACGCCAGGGGCTTCCGAGGTGTTCACA 1304

Qy      879 ----------------------------GACTCCTTCCTCCGATTTTTGGAAAGAAGTGC 910
```

```
                                                    ||||||||||||||||||||||||||||||||
Db       1305 CCTGCAAACCGCCCCGACCTGGCCCCCAGACTCCTTCCTCCGATTTTTGGAAAGAAGTGC 1364

Qy        911 AGCTGGCGCTGCCGCGGAAGGTGCGGTACAGGCCGGTGGAAGGAGACCCCCAGACGCAGC 970
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1365 AGCTGGCGCTGCCGCGGAAGGTGCGGTACAGGCCGGTGGAAGGAGACCCCCAGACGCAGC 1424

Qy        971 TGCAGGACGACAAGGACCCCATGCTGATTCTTCGAGGCCGAGTCCCTGAGGGCCGGGCCC 1030
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1425 TGCAGGACGACAAGGACCCCATGCTGATTCTTCGAGGCCGAGTCCCTGAGGGCCGGGCCC 1484

Qy       1031 TGGACTCAGAGGTGGACCCGGACCCTGAGGGCGACCTGGGTGTCCGGGGGCCTGTTTTTG 1090
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1485 TGGACTCAGAGGTGGACCCGGACCCTGAGGGCGACCTGGGTGTCCGGGGGCCTGTTTTTG 1544

Qy       1091 GAGAGCCATCAGCTCCACCGCACACCAGTGGGGTCTCGCTGGGAGAGAGCCGGAGCAGCG 1150
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1545 GAGAGCCATCAGCTCCACCGCACACCAGTGGGGTCTCGCTGGGAGAGAGCCGGAGCAGCG 1604

Qy       1151 AAGTGGACGTCTCGGATCTCGGCTCGCGAAACTACAGTGCCCGCACAGACTTCTACTGCC 1210
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1605 AAGTGGACGTCTCGGATCTCGGCTCGCGAAACTACAGTGCCCGCACAGACTTCTACTGCC 1664

Qy       1211 TGGTGTCCAAGGATGATATGTAG 1233
              |||||||||||||||||||||||
Db       1665 TGGTGTCCAAGGATGATATGTAG 1687
```

```
RESULT 14
BV177172/c
LOCUS       BV177172                655 bp    DNA     linear   STS 10-JUN-2004
DEFINITION  sqnm93226 Human DNA (Sequenom) Homo sapiens STS genomic, sequence
            tagged site.
ACCESSION   BV177172
VERSION     BV177172.1  GI:48013190
KEYWORDS    STS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 655)
  AUTHORS   Nelson,R.M., Marnellos,G., Kammerer,S., Hoyal,C.R., Shi,M.M.,
            Cantor,C.R. and Braun,A.
  TITLE     Large-Scale Validation of Single Nucleotide Polymorphisms in Gene
            Regions
  JOURNAL   Genome Res. (2004) In press
COMMENT
            Contact: Andreas Braun
            Pharmaceuticals division
            Sequenom, Inc.
            3595 John Hopkins Court, San Diego, CA 92121, USA
            Tel: 18582029018
            Fax: 18582029020
            Email: abraun@sequenom.com
            Primer A: No primer sequence submitted
            Primer B: No primer sequence submitted
            STS size: 655.
FEATURES             Location/Qualifiers
     source          1. .655
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone_lib="Human DNA (Sequenom)"
     STS             <1. .>655
ORIGIN

  Query Match           32.3%;  Score 398;  DB 11;  Length 655;
  Best Local Similarity   97.8%;  Pred. No. 2.7e-52;
  Matches  402;  Conservative    1;  Mismatches    8;  Indels    0;  Gaps    0;

Qy        823 CTGCTGCGCCAGCACCGCCACCTGGTGACCTTGCTGCTCTGGAGGCCCGGCTCCGTGACT 882
              | |||||  |   ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        611 CCGCTGCTGCGCAACCGCCACCTGGTGACCTTGCTGCTCTGGAGGCCCGGCTCCGTGACT 552

Qy        883 CCTTCCTCCGATTTTTGGAAAGAAGTGCAGCTGGCGCTGCCGCGGAAGGTGCGGTACAGG 942
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db        551 CCTTCCTCCGATTTTTGGAAAGAAGTGCAGCTGGCGCTGCCGCGGAAGGTGCAGTACAGG 492

Qy        943 CCGGTGGAAGGAGACCCCCAGACGCAGCTGCAGGACGACAAGGACCCCATGCTGATTCTT 1002
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        491 CCTGTGGAAGGAGACCCCCAGACGCAGCTGCAGGACGACAAGGACCCCATGCTGATTCTT 432

Qy       1003 CGAGGCCGAGTCCCTGAGGGCCGGGCCCTGGACTCAGAGGTGGACCCGGACCCTGAGGGC 1062
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        431 CGAGGCCGAGTCCCTGAGGGCCGGGCCCTGGACTCAGAGGTGGACCCGGACCCTGAGGGC 372

Qy       1063 GACCTGGGTGTCCGGGGGCCTGTTTTTGGAGAGCCATCAGCTCCACCGCACACCAGTGGG 1122
              |||||||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||
Db        371 GACCTGGGTGTCCGGGGGCCTGTYTTTGGAGAGCCATCAGCTCCACCGCACACCAGTGGG 312

Qy       1123 GTCTCGCTGGGAGAGAGCCGGAGCAGCGAAGTGGACGTCTCGGATCTCGGCTCGCGAAAC 1182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        311 GTCTCGCTGGGAGAGAGCCGGAGCAGCGAAGTGGACGTCTCGGATCTCGGCTCGCGAAAC 252

Qy       1183 TACAGTGCCCGCACAGACTTCTACTGCCTGGTGTCCAAGGATGATATGTAG 1233
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        251 TACAGTGCCCGCACAGACTTCTACTGCCTGGTGTCCAAGGATGATATGTAG 201
```

```
RESULT 15
AX525866
LOCUS       AX525866                641 bp    DNA     linear   PAT 22-NOV-2002
DEFINITION  Sequence 388 from Patent WO02066682.
ACCESSION   AX525866
VERSION     AX525866.1  GI:25170744
KEYWORDS    .
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1
  AUTHORS   Farris,G., Hicken,S.H. and Farr,S.B.
  TITLE     Rat toxicologically relevant genes and uses thereof
  JOURNAL   Patent: WO 02066682-A 388 29-AUG-2002;
            Phase-1 Molecular Toxicology Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .641
                     /organism="Rattus norvegicus"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:10116"
ORIGIN

  Query Match           31.8%;  Score 391.6;  DB 6;  Length 641;
  Best Local Similarity   81.4%;  Pred. No. 2.6e-51;
  Matches  454;  Conservative    0;  Mismatches  104;  Indels    0;  Gaps    0;

Qy        673 ATCGTGGTGCTTTCGGACGCCTTCCTGAGCCGGGCCTGGTGCAGCCACAGCTTCCGGGAG 732
              |||||||| ||||| || |||||||| || ||  ||||||| ||||| ||||||||||||
Db         84 ATCGTGGTTCTTTCAGATGCCTTCCTCAGACGTCCCTGGTGTAGCCAGAGCTTCCGGGAG 143

Qy        733 GGCCTGTGCCGGCTGCTGGAGCTCACCCGCAGACCCATCTTCATCACCTTCGAGGGCCAG 792
              || ||||||||| || |||||||||||||||||||| ||||| |||||||||||||||||||
Db        144 GGACTGTGCCGCCTACTGGAGCTCACCCGCAGACCTATCTTTATCACCTTCGAGGGCCAG 203

Qy        793 AGGCGCGACCCCGCGCACCCGGCGCTCCGCCTGCTGCGCCAGCACCGCCACCTGGTGACC 852
              ||||| || |||   ||||| || ||||| || |||||||||||||||||||| ||||||
Db        204 AGGCGTGAGCCCATTCACCCTGCTCTCCGTCTCCTGCGCCAGCACCGCCACCTCGTGACC 263

Qy        853 TTGCTGCTCTGGAGGCCCGGCTCCGTGACTCCTTCCTCCGATTTTTGGAAAGAAGTGCAG 912
              ||||||| |||| ||| |||||||||| |||||||| |||||||||||||  | |||
Db        264 CTGCTGCTTTGGAAGCCTGGCTCCGTGACACCTTCCTCTGATTTTTGGAAAGAGCTACAG 323

Qy        913 CTGGCGCTGCCGCGGAAGGTGCGGTACAGGCCGGTGGAAGGAGACCCCCAGACGCAGCTG 972
              |||| || || ||||||||||  ||||| || ||||| |||||||||| || |  ||
Db        324 TTGGCACTACCACGGAAGGTGCAATACAGACCCGTGGAGGGAGACCCCCAAACCCGACTT 383
```

```
Qy        973 CAGGACGACAAGGACCCCATGCTGATTCTTCGAGGCCGAGTCCCTGAGGGCCGGGCCCTG 1032
              ||||| ||||| || |||||||| ||   |  |||| || |    |  |||||||||| ||
Db        384 CAGGATGACAAAGATCCCATGCTAATCGTGAGAGGACGTGCTGCCCAGGGCCGGGCCATG 443

Qy       1033 GACTCAGAGGTGGACCCGGACCCTGAGGGCGACCTGGGTGTCCGGGGGCCTGTTTTTGGA 1092
              || |||||| |||| || |||||||||||| |||||||||||||| || ||||| |||||
Db        444 GAGTCAGAGCTGGATCCAGACCCTGAGGGAGACCTGGGTGTCCGTGGACCTGTCTTTGGG 503

Qy       1093 GAGCCATCAGCTCCACCGCACACCAGTGGGGTCTCGCTGGGAGAGAGCCGGAGCAGCGAA 1152
              |||||| || |||||| ||| || |||||| | |||||| ||| ||| |||
Db        504 GAGCCACCAACTCCACTACACGAAAGCAAGGTCTCCATAGGAGAGGGCCATGCCAGTGAA 563

Qy       1153 GTGGACGTCTCGGATCTCGGCTCGCGAAACTACAGTGCCCGCACAGACTTCTACTGCCTG 1212
               |||| ||||| || || ||||| |||||||||||||| || ||||||||||||||||
Db        564 ATGGATGTCTCTGACCTTGGCTCTCGAAACTACAGTGCGCGAACAGACTTCTACTGCCTT 623

Qy       1213 GTGTCCAAGGATGATATG 1230
              |||||  |||||||  ||
Db        624 GTGTCTGAGGATGACGTG 641
```

Search completed: July 30, 2005, 00:04:23
Job time : 5666 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: July 29, 2005, 20:10:25 ; Search time 757 Seconds
(without alignments)
9642.062 Million cell updates/sec

Title: US-09-598-443-1
Perfect score: 1233
Sequence: 1 atgccaggtgtctgtgatag.....tgtccaaggatgatatgtag 1233

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1233	100.0	1233	2	AAX88091	Aax88091 Human SIG
2	1233	100.0	1694	13	ADP54077	Adp54077 Human PRO
3	1233	100.0	1695	12	ADO20094	Ado20094 Human PRO
4	1233	100.0	1695	12	ADO19086	Ado19086 Human PRO
5	1233	100.0	1695	13	ADP22969	Adp22969 PRO polyp
6	1231.4	99.9	1659	4	AAF27859	Aaf27859 Human NOV
7	1231.4	99.9	1659	9	ACD40262	Acd40262 Human bre
8	1231.4	99.9	1659	11	ADM56382	Adm56382 Human cDN
9	1231.4	99.9	1659	12	ADF66744	Adf66744 cDNA enco
10	1231.4	99.9	1659	12	ADO60254	Ado60254 Human NOV
11	1230.4	99.8	1659	12	ADI19781	Adi19781 Human NOV
12	1156	93.8	1833	12	ADP46448	Adp46448 Human col
13	1071.8	86.9	1649	6	ABK40277	Abk40277 cDNA enco
14	1071.8	86.9	1649	10	ADJ37328	Adj37328 Human tum
15	1071.8	86.9	1649	12	ADG68252	Adg68252 Human PRO
16	1029.4	83.5	1428	4	AAS27325	Aas27325 cDNA enco
17	1029.4	83.5	1428	4	AAI63958	Aai63958 Human pol
18	1029.4	83.5	1428	10	ADB93503	Adb93503 Human cDN
19	1029.4	83.5	1428	12	ADM24509	Adm24509 Human PRO
20	1029.4	83.5	1442	4	AAI63889	Aai63889 Human pol
21	1029.4	83.5	1442	12	ADM24440	Adm24440 Human PRO
22	1025	83.1	1296	3	AAF18292	Aaf18292 Lung canc
23	846.6	68.7	1090	7	ADR41330	Adr41330 Human CD-
24	841	68.2	2138	12	ADP46449	Adp46449 Human col
c 25	506.4	41.1	851	4	AAH99795	Aah99795 Human pro
c 26	391.8	31.8	575	13	ADR26100	Adr26100 Breast ca
27	391.6	31.8	641	6	ABT09300	Abt09300 Phase-1 R
28	391.6	31.8	641	10	ADG31018	Adg31018 Liver tox
29	391.6	31.8	641	13	ADR91262	Adr91262 Spleen ne
30	357	29.0	676	12	ADP46450	Adp46450 Human col
c 31	250.6	20.3	557	6	ABK62170	Abk62170 Rat seque
c 32	250.6	20.3	557	10	ADB55156	Adb55156 Toxicity-
c 33	250.6	20.3	557	10	ADB49677	Adb49677 Primary r
34	245.4	19.9	249	3	AAC06025	Aac06025 Human sec
35	231.8	18.8	322	12	ADL87643	Adl87643 DNA up-re
36	231.8	18.8	322	12	ADL87644	Adl87644 DNA up-re
37	196.8	16.0	407	12	ADL87641	Adl87641 DNA up-re
38	196.8	16.0	407	12	ADL87642	Adl87642 DNA up-re
c 39	186	15.1	388	6	ABN94880	Abn94880 Gene #137
40	161.6	13.1	266	4	AAK61892	Aak61892 Human imm
c 41	69.8	5.7	259	4	AAK53693	Aak53693 Murine tr
42	65.4	5.3	567	13	ADR05988	Adr05988 Human mut
43	62.2	5.0	567	13	ADR05992	Adr05992 Human mut
44	60.8	4.9	570	3	ABL50864	Abl50864 Modified
45	60.6	4.9	567	13	ADR05990	Adr05990 Human mut

ALIGNMENTS

RESULT 1
AAX88091
ID AAX88091 standard; DNA; 1233 BP.
XX
AC AAX88091;
XX
DT 08-SEP-1999 (first entry)
XX
DE Human SIGIRR DNA.
XX
KW SIGIRR; anti-inflammatory; anti-autoimmune disease; chromosome 11;
KW interleukin-1 receptor; 11p15.5; arthrogryposis multiplex congenita;
KW breast cancer; insulin-dependent diabetes; sickle cell anaemia;
KW bladder cancer; detection; cell-signal transduction; gene therapy;
KW immune regulation; cell proliferation; cell death; cell migration;
KW cell interaction; inflammation; marker; mass spectrometry; human; ss.
XX
OS Homo sapiens.
XX
PN WO9932626-A1.
XX
PD 01-JUL-1999.
XX
PF 23-DEC-1998; 98WO-US027368.
XX
PR 23-DEC-1997; 97US-0068770P.
XX
PA (IMMV) IMMUNEX CORP.
XX
PI Sims JE;
XX
DR WPI; 1999-418925/35.
DR P-PSDB; AAY25426.
XX
PT New interleukin-1 receptor analog SIGIRR nucleic acid and proteins.
XX
PS Claim 1; Page 68-69; 72pp; English.
XX
CC This invention describes a novel human SIGIRR DNA, its allelic variants
CC or species homologs which have anti-inflammatory and anti-autoimmune
CC disease activity. SIGIRR is an analog of interleukin-1 receptor. The
CC products of the invention are used (a) as probes or primers for
CC identifying nucleic acid that encodes proteins with SIGIRR activity; (b)

```
CC   to identify human chromosome 11, to map genes on this chromosome and to
CC   identify disease-related genes (particularly in the region 11p15.5 where
CC   genes are present associated with e.g. arthrogryposis multiplex
CC   congenita, breast cancer, insulin-dependent diabetes, sickle cell
CC   anaemia, bladder cancer), including detection of defective genes; (c) to
CC   study cell-signal transduction and the SIGIRR system, and (d) in gene
CC   therapy. Sense and antisense oligonucleotides can be used to inhibit
CC   expression of the SIGIRR gene. The proteins of the invention are used:
CC   (i) to study cellular processes (immune regulation, proliferation, death,
CC   migration, interaction with other cells and inflammation); (ii) to
CC   identify and purify proteins that associate with SIGIRR ligands and
CC   receptors, and to measure their biological activity; (iii) in screening
CC   for, and rational design of, potential inhibitors of activity; (iv)
CC   therapeutically against diseases mediated by SIGIRR polypeptide counter-
CC   structures; (v) as molecular weight (m.w.) markers in electrophoresis;
CC   (vi) for determining isoelectric points of unknown proteins; (vii) as
CC   controls for determining the extent of protein fragmentation (e.g. to aid
CC   characterization of protein structures by mass spectrometry); (viii) for
CC   generation of antibodies (Ab); (ix) to deliver diagnostic or therapeutic
CC   agents to cells that express SIGIRR binding molecules
XX
SQ   Sequence 1233 BP; 203 A; 408 C; 380 G; 242 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 1233;  DB 2;  Length 1233;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-236;
  Matches 1233;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGCCAGGTGTCTGTGATAGGGCCCCTGACTTCCTCTCCCCGTCTGAAGACCAGGTGCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCCAGGTGTCTGTGATAGGGCCCCTGACTTCCTCTCCCCGTCTGAAGACCAGGTGCTG 60

Qy         61 AGGCCTGCCTTGGGCAGCTCAGTGGCTCTGAACTGCACGGCTTGGGTAGTCTCTGGGCCC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGGCCTGCCTTGGGCAGCTCAGTGGCTCTGAACTGCACGGCTTGGGTAGTCTCTGGGCCC 120

Qy        121 CACTGCTCCCTGCCTTCAGTCCAGTGGCTGAAAGACGGGCTTCCATTGGGAATTGGGGGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CACTGCTCCCTGCCTTCAGTCCAGTGGCTGAAAGACGGGCTTCCATTGGGAATTGGGGGC 180

Qy        181 CACTACAGCCTCCACGAGTACTCCTGGGTCAAGGCCAACCTGTCAGAGGTGCTTGTGTCC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CACTACAGCCTCCACGAGTACTCCTGGGTCAAGGCCAACCTGTCAGAGGTGCTTGTGTCC 240

Qy        241 AGTGTCCTGGGGGTCAACGTGACCAGCACTGAAGTCTATGGGGCCTTCACCTGCTCCATC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGTGTCCTGGGGGTCAACGTGACCAGCACTGAAGTCTATGGGGCCTTCACCTGCTCCATC 300

Qy        301 CAGAACATCAGCTTCTCCTCCTTCACTCTTCAGAGAGCTGGCCCTACAAGCCACGTGGCT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CAGAACATCAGCTTCTCCTCCTTCACTCTTCAGAGAGCTGGCCCTACAAGCCACGTGGCT 360

Qy        361 GCGGTGCTGGCCTCCCTCCTGGTCCTGCTGGCCCTGCTGCTGGCCGCCCTGCTCTATGTC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCGGTGCTGGCCTCCCTCCTGGTCCTGCTGGCCCTGCTGCTGGCCGCCCTGCTCTATGTC 420

Qy        421 AAGTGCCGTCTCAACGTGCTGCTCTGGTACCAGGACGCGTATGGGGAGGTGGAGATAAAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AAGTGCCGTCTCAACGTGCTGCTCTGGTACCAGGACGCGTATGGGGAGGTGGAGATAAAC 480

Qy        481 GACGGGAAGCTCTACGACGCCTACGTCTCCTACAGCGACTGCCCCGAGGACCGCAAGTTC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GACGGGAAGCTCTACGACGCCTACGTCTCCTACAGCGACTGCCCCGAGGACCGCAAGTTC 540

Qy        541 GTGAACTTCATCCTAAAGCCGCAGCTGGAGCGGCGTCGGGGCTACAAGCTCTTCCTGGAC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GTGAACTTCATCCTAAAGCCGCAGCTGGAGCGGCGTCGGGGCTACAAGCTCTTCCTGGAC 600

Qy        601 GACCGCGACCTCCTGCCGCGCGCTGAGCCCTCCGCCGACCTCTTGGTGAACCTGAGCCGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GACCGCGACCTCCTGCCGCGCGCTGAGCCCTCCGCCGACCTCTTGGTGAACCTGAGCCGC 660

Qy        661 TGCCGACGCCTCATCGTGGTGCTTTCGGACGCCTTCCTGAGCCGGGCCTGGTGCAGCCAC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TGCCGACGCCTCATCGTGGTGCTTTCGGACGCCTTCCTGAGCCGGGCCTGGTGCAGCCAC 720

Qy        721 AGCTTCCGGGAGGGCCTGTGCCGGCTGCTGGAGCTCACCCGCAGACCCATCTTCATCACC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AGCTTCCGGGAGGGCCTGTGCCGGCTGCTGGAGCTCACCCGCAGACCCATCTTCATCACC 780

Qy        781 TTCGAGGGCCAGAGGCGCGACCCCGCGCACCCGGCGCTCCGCCTGCTGCGCCAGCACCGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TTCGAGGGCCAGAGGCGCGACCCCGCGCACCCGGCGCTCCGCCTGCTGCGCCAGCACCGC 840

Qy        841 CACCTGGTGACCTTGCTGCTCTGGAGGCCCGGCTCCGTGACTCCTTCCTCCGATTTTTGG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CACCTGGTGACCTTGCTGCTCTGGAGGCCCGGCTCCGTGACTCCTTCCTCCGATTTTTGG 900

Qy        901 AAAGAAGTGCAGCTGGCGCTGCCGCGGAAGGTGCGGTACAGGCCGGTGGAAGGAGACCCC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AAAGAAGTGCAGCTGGCGCTGCCGCGGAAGGTGCGGTACAGGCCGGTGGAAGGAGACCCC 960

Qy        961 CAGACGCAGCTGCAGGACGACAAGGACCCCATGCTGATTCTTCGAGGCCGAGTCCCTGAG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 CAGACGCAGCTGCAGGACGACAAGGACCCCATGCTGATTCTTCGAGGCCGAGTCCCTGAG 1020

Qy       1021 GGCCGGGCCCTGGACTCAGAGGTGGACCCGGACCCTGAGGGCGACCTGGGTGTCCGGGGG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GGCCGGGCCCTGGACTCAGAGGTGGACCCGGACCCTGAGGGCGACCTGGGTGTCCGGGGG 1080

Qy       1081 CCTGTTTTTGGAGAGCCATCAGCTCCACCGCACACCAGTGGGGTCTCGCTGGGAGAGAGC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CCTGTTTTTGGAGAGCCATCAGCTCCACCGCACACCAGTGGGGTCTCGCTGGGAGAGAGC 1140

Qy       1141 CGGAGCAGCGAAGTGGACGTCTCGGATCTCGGCTCGCGAAACTACAGTGCCCGCACAGAC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CGGAGCAGCGAAGTGGACGTCTCGGATCTCGGCTCGCGAAACTACAGTGCCCGCACAGAC 1200

Qy       1201 TTCTACTGCCTGGTGTCCAAGGATGATATGTAG 1233
              |||||||||||||||||||||||||||||||||
Db       1201 TTCTACTGCCTGGTGTCCAAGGATGATATGTAG 1233


RESULT 2
ADP54077
ID   ADP54077 standard; cDNA; 1694 BP.
XX
AC   ADP54077;
XX
DT   18-NOV-2004  (first entry)
XX
DE   Human PRO cDNA sequence SEQ ID NO:53.
XX
KW   human; PRO; immune related disease; inflammatory immune response;
KW   immune response stimulation; antiallergic; antianaemic; antiarthritic;
KW   antiasthmatic; antidiabetic; antiinflammatory; antipsoriatic;
KW   antirheumatic; antithyroid; CNS; dermatological; gastrointestinal;
KW   haemostatic; hepatotropic; immunostimulant; immunosuppressive; muscular;
KW   nephrotropic; neuroprotective; osteopathic; respiratory; vasotropic;
KW   virucide; gene therapy; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO2004039956-A2.
XX
PD   13-MAY-2004.
XX
PF   28-OCT-2003; 2003WO-US034381.
XX
PR   29-OCT-2002; 2002US-0422472P.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Aggarwal S,  Clark H,  Gurney AL,  Schoenfeld J,  Williams PM;
PI   Wood WI,  Wu TD;
XX
DR   WPI; 2004-376182/35.
```

```
DR   P-PSDB; ADP54078.
XX
PT   New PRO polynucleotides and polypeptides, useful in useful in diagnosing
PT   and treating an immune related disease, e.g. systemic lupus
PT   erythematosus, rheumatoid arthritis, diabetes mellitus or asthma and in
PT   stimulating an immune response.
XX
PS   Claim 2; SEQ ID NO 53; 3009pp; English.
XX
CC   The present invention describes an isolated PRO nucleic acid (I). Also
CC   described: (1) a vector comprising (I); (2) a host cell comprising the
CC   vector of (1); (3) a process for producing a PRO polypeptides; (4) an
CC   isolated PRO polypeptide; (5) a chimeric molecule comprising the
CC   polypeptide of (4) fused to a heterologous amino acid sequence; (6) an
CC   antibody which specifically binds to a polypeptide of (4); (7) a
CC   composition of matter comprising a polypeptide of (4), an agonist or
CC   antagonist of the polypeptide or an antibody that binds to the
CC   polypeptide in combination with a carrier; (8) an article of manufacture
CC   comprising a container, a label on the container and a composition of
CC   matter of (7); (9) a method of treating an immune related disease in a
CC   mammal; (10) a method for determining the presence of a PRO polypeptide
CC   in a sample suspected of having the polypeptide; (11) a method of
CC   diagnosing an immune related disease or an inflammatory immune response
CC   in mammal; (12) a method of identifying a compound that inhibits or
CC   mimics the activity of or expression of a gene encoding a PRO polypeptide
CC   ; and (13) a method of stimulating the immune response in a mammal. The
CC   PRO sequences have antiallergic, antianaemic, antiarthritic,
CC   antiasthmatic, antidiabetic, antiinflammatory, antipsoriatic,
CC   antirheumatic, antithyroid, CNS, dermatological, gastrointestinal,
CC   haemostatic, hepatotropic, immunostimulant, immunosuppressive, muscular,
CC   nephrotropic, neuroprotective, osteopathic, respiratory, vasotropic and
CC   virucide activities, and can be used in gene therapy. The nucleic acid
CC   (I) and the encoded polypeptides, compositions, kits and methods are
CC   useful in diagnosing and treating an immune related disease and in
CC   stimulating an immune response. The present sequence represents a human
CC   PRO nucleotide sequence from the present invention.
XX
SQ   Sequence 1694 BP; 328 A; 556 C; 510 G; 300 T; 0 U; 0 Other;

  Query Match           100.0%;  Score 1233;  DB 13;  Length 1694;
  Best Local Similarity  100.0%;  Pred. No. 1.1e-236;
  Matches 1233;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGCCAGGTGTCTGTGATAGGGCCCCTGACTTCCTCTCCCCGTCTGAAGACCAGGTGCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 ATGCCAGGTGTCTGTGATAGGGCCCCTGACTTCCTCTCCCCGTCTGAAGACCAGGTGCTG 286

Qy         61 AGGCCTGCCTTGGGCAGCTCAGTGGCTCTGAACTGCACGGCTTGGGTAGTCTCTGGGCCC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 AGGCCTGCCTTGGGCAGCTCAGTGGCTCTGAACTGCACGGCTTGGGTAGTCTCTGGGCCC 346

Qy        121 CACTGCTCCCTGCCTTCAGTCCAGTGGCTGAAAGACGGGCTTCCATTGGGAATTGGGGGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        347 CACTGCTCCCTGCCTTCAGTCCAGTGGCTGAAAGACGGGCTTCCATTGGGAATTGGGGGC 406

Qy        181 CACTACAGCCTCCACGAGTACTCCTGGGTCAAGGCCAACCTGTCAGAGGTGCTTGTGTCC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        407 CACTACAGCCTCCACGAGTACTCCTGGGTCAAGGCCAACCTGTCAGAGGTGCTTGTGTCC 466

Qy        241 AGTGTCCTGGGGGTCAACGTGACCAGCACTGAAGTCTATGGGGCCTTCACCTGCTCCATC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        467 AGTGTCCTGGGGGTCAACGTGACCAGCACTGAAGTCTATGGGGCCTTCACCTGCTCCATC 526

Qy        301 CAGAACATCAGCTTCTCCTCCTTCACTCTTCAGAGAGCTGGCCCTACAAGCCACGTGGCT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        527 CAGAACATCAGCTTCTCCTCCTTCACTCTTCAGAGAGCTGGCCCTACAAGCCACGTGGCT 586

Qy        361 GCGGTGCTGGCCTCCCTCCTGGTCCTGCTGGCCCTGCTGCTGGCCGCCCTGCTCTATGTC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        587 GCGGTGCTGGCCTCCCTCCTGGTCCTGCTGGCCCTGCTGCTGGCCGCCCTGCTCTATGTC 646

Qy        421 AAGTGCCGTCTCAACGTGCTGCTCTGGTACCAGGACGCGTATGGGGAGGTGGAGATAAAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        647 AAGTGCCGTCTCAACGTGCTGCTCTGGTACCAGGACGCGTATGGGGAGGTGGAGATAAAC 706

Qy        481 GACGGGAAGCTCTACGACGCCTACGTCTCCTACAGCGACTGCCCCGAGGACCGCAAGTTC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        707 GACGGGAAGCTCTACGACGCCTACGTCTCCTACAGCGACTGCCCCGAGGACCGCAAGTTC 766

Qy        541 GTGAACTTCATCCTAAAGCCGCAGCTGGAGCGGCGTCGGGGCTACAAGCTCTTCCTGGAC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        767 GTGAACTTCATCCTAAAGCCGCAGCTGGAGCGGCGTCGGGGCTACAAGCTCTTCCTGGAC 826

Qy        601 GACCGCGACCTCCTGCCGCGCGCTGAGCCCTCCGCCGACCTCTTGGTGAACCTGAGCCGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        827 GACCGCGACCTCCTGCCGCGCGCTGAGCCCTCCGCCGACCTCTTGGTGAACCTGAGCCGC 886

Qy        661 TGCCGACGCCTCATCGTGGTGCTTTCGGACGCCTTCCTGAGCCGGGCCTGGTGCAGCCAC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        887 TGCCGACGCCTCATCGTGGTGCTTTCGGACGCCTTCCTGAGCCGGGCCTGGTGCAGCCAC 946

Qy        721 AGCTTCCGGGAGGGCCTGTGCCGGCTGCTGGAGCTCACCCGCAGACCCATCTTCATCACC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        947 AGCTTCCGGGAGGGCCTGTGCCGGCTGCTGGAGCTCACCCGCAGACCCATCTTCATCACC 1006

Qy        781 TTCGAGGGCCAGAGGCGCGACCCCGCGCACCCGGCGCTCCGCCTGCTGCGCCAGCACCGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1007 TTCGAGGGCCAGAGGCGCGACCCCGCGCACCCGGCGCTCCGCCTGCTGCGCCAGCACCGC 1066

Qy        841 CACCTGGTGACCTTGCTGCTCTGGAGGCCCGGCTCCGTGACTCCTTCCTCCGATTTTTGG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1067 CACCTGGTGACCTTGCTGCTCTGGAGGCCCGGCTCCGTGACTCCTTCCTCCGATTTTTGG 1126

Qy        901 AAAGAAGTGCAGCTGGCGCTGCCGCGGAAGGTGCGGTACAGGCCGGTGGAAGGAGACCCC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1127 AAAGAAGTGCAGCTGGCGCTGCCGCGGAAGGTGCGGTACAGGCCGGTGGAAGGAGACCCC 1186

Qy        961 CAGACGCAGCTGCAGGACGACAAGGACCCCATGCTGATTCTTCGAGGCCGAGTCCCTGAG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1187 CAGACGCAGCTGCAGGACGACAAGGACCCCATGCTGATTCTTCGAGGCCGAGTCCCTGAG 1246

Qy       1021 GGCCGGGCCCTGGACTCAGAGGTGGACCCGGACCCTGAGGGCGACCTGGGTGTCCGGGGG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1247 GGCCGGGCCCTGGACTCAGAGGTGGACCCGGACCCTGAGGGCGACCTGGGTGTCCGGGGG 1306

Qy       1081 CCTGTTTTTGGAGAGCCATCAGCTCCACCGCACACCAGTGGGGTCTCGCTGGGAGAGAGC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1307 CCTGTTTTTGGAGAGCCATCAGCTCCACCGCACACCAGTGGGGTCTCGCTGGGAGAGAGC 1366

Qy       1141 CGGAGCAGCGAAGTGGACGTCTCGGATCTCGGCTCGCGAAACTACAGTGCCCGCACAGAC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1367 CGGAGCAGCGAAGTGGACGTCTCGGATCTCGGCTCGCGAAACTACAGTGCCCGCACAGAC 1426

Qy       1201 TTCTACTGCCTGGTGTCCAAGGATGATATGTAG 1233
              ||||||||||||||||||||||||||||||||
Db       1427 TTCTACTGCCTGGTGTCCAAGGATGATATGTAG 1459


RESULT 3
ADO20094
ID   ADO20094 standard; cDNA; 1695 BP.
XX
AC   ADO20094;
XX
DT   12-AUG-2004  (first entry)
XX
DE   Human PRO polynucleotide #502.
XX
KW   Human; PRO; gene; ss; immune related disorder;
KW   systemic lupus erythematosus; rheumatoid arthritis; osteoarthritis;
KW   juvenile chronic arthritis; systemic sclerosis; Sjogren's syndrome;
KW   vasculitis; sarcoidosis; autoimmune haemolytic anaemia;
KW   autoimmune thrombocytopenia; thyroiditis; diabetes mellitus;
KW   renal disease; demyelinating disease; central nervous system;
KW   peripheral nervous system; demyelinating polyneuropathy;
KW   Guillain-Barre syndrome;
```

```
KW   chronic inflammatory demyelinating polyneuropathy.
XX
OS   Homo sapiens.
XX
PN   WO2004043361-A2.
XX
PD   27-MAY-2004.
XX
PF   06-NOV-2003; 2003WO-US035268.
XX
PR   08-NOV-2002; 2002US-0425235P.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Fong S,  Dennis K,  Clark H,  Chiu H,  Schoenfeld J,  Williams PM;
PI   Wood WI,  Wu TD;
XX
DR   WPI; 2004-420067/39.
DR   P-PSDB; ADO20095.
XX
PT   Novel PRO polypeptide e.g., PRO69614, PRO71106, or PRO86388 useful for
PT   treating an immune related disorder such as systemic lupus erythematosus,
PT   rheumatoid arthritis, osteoarthritis, juvenile chronic arthritis or
PT   spondyloarthropathy.
XX
PS   Claim 1; SEQ ID NO 1058; 1731pp; English.
XX
CC   The invention relates to human PRO polypeptides and the polynucleotides
CC   encoding them. The polypeptides and polynucleotides are useful for
CC   treating and diagnosing immune related disorders in mammals. The immune
CC   related disorders include systemic lupus erythematosus, rheumatoid
CC   arthritis, osteoarthritis, juvenile chronic arthritis, systemic
CC   sclerosis, Sjogren's syndrome, vasculitis, sarcoidosis, autoimmune
CC   haemolytic anaemia, autoimmune thrombocytopenia, thyroiditis, diabetes
CC   mellitus, immune-mediated renal disease, demyelinating diseases of the
CC   central or peripheral nervous system, demyelinating polyneuropathy,
CC   Guillain-Barre syndrome and chronic inflammatory demyelinating
CC   polyneuropathy. This sequence represents a human PRO polynucleotide of
CC   the invention.
XX
SQ   Sequence 1695 BP; 329 A; 556 C; 510 G; 300 T; 0 U; 0 Other;

  Query Match            100.0%;  Score 1233;  DB 12;  Length 1695;
  Best Local Similarity  100.0%;  Pred. No. 1.1e-236;
  Matches 1233;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGCCAGGTGTCTGTGATAGGGCCCCTGACTTCCTCTCCCCGTCTGAAGACCAGGTGCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 ATGCCAGGTGTCTGTGATAGGGCCCCTGACTTCCTCTCCCCGTCTGAAGACCAGGTGCTG 286

Qy         61 AGGCCTGCCTTGGGCAGCTCAGTGGCTCTGAACTGCACGGCTTGGGTAGTCTCTGGGCCC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 AGGCCTGCCTTGGGCAGCTCAGTGGCTCTGAACTGCACGGCTTGGGTAGTCTCTGGGCCC 346

Qy        121 CACTGCTCCCTGCCTTCAGTCCAGTGGCTGAAAGACGGGCTTCCATTGGGAATTGGGGGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        347 CACTGCTCCCTGCCTTCAGTCCAGTGGCTGAAAGACGGGCTTCCATTGGGAATTGGGGGC 406

Qy        181 CACTACAGCCTCCACGAGTACTCCTGGGTCAAGGCCAACCTGTCAGAGGTGCTTGTGTCC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        407 CACTACAGCCTCCACGAGTACTCCTGGGTCAAGGCCAACCTGTCAGAGGTGCTTGTGTCC 466

Qy        241 AGTGTCCTGGGGGTCAACGTGACCAGCACTGAAGTCTATGGGGCCTTCACCTGCTCCATC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        467 AGTGTCCTGGGGGTCAACGTGACCAGCACTGAAGTCTATGGGGCCTTCACCTGCTCCATC 526

Qy        301 CAGAACATCAGCTTCTCCTCCTTCACTCTTCAGAGAGCTGGCCCTACAAGCCACGTGGCT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        527 CAGAACATCAGCTTCTCCTCCTTCACTCTTCAGAGAGCTGGCCCTACAAGCCACGTGGCT 586

Qy        361 GCGGTGCTGGCCTCCCTCCTGGTCCTGCTGGCCCTGCTGCTGGCCGCCCTGCTCTATGTC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        587 GCGGTGCTGGCCTCCCTCCTGGTCCTGCTGGCCCTGCTGCTGGCCGCCCTGCTCTATGTC 646

Qy        421 AAGTGCCGTCTCAACGTGCTGCTCTGGTACCAGGACGCGTATGGGGAGGTGGAGATAAAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        647 AAGTGCCGTCTCAACGTGCTGCTCTGGTACCAGGACGCGTATGGGGAGGTGGAGATAAAC 706

Qy        481 GACGGGAAGCTCTACGACGCCTACGTCTCCTACAGCGACTGCCCCGAGGACCGCAAGTTC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        707 GACGGGAAGCTCTACGACGCCTACGTCTCCTACAGCGACTGCCCCGAGGACCGCAAGTTC 766

Qy        541 GTGAACTTCATCCTAAAGCCGCAGCTGGAGCGGCGTCGGGGCTACAAGCTCTTCCTGGAC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        767 GTGAACTTCATCCTAAAGCCGCAGCTGGAGCGGCGTCGGGGCTACAAGCTCTTCCTGGAC 826

Qy        601 GACCGCGACCTCCTGCCGCGCGCTGAGCCCTCCGCCGACCTCTTGGTGAACCTGAGCCGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        827 GACCGCGACCTCCTGCCGCGCGCTGAGCCCTCCGCCGACCTCTTGGTGAACCTGAGCCGC 886

Qy        661 TGCCGACGCCTCATCGTGGTGCTTTCGGACGCCTTCCTGAGCCGGGCCTGGTGCAGCCAC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        887 TGCCGACGCCTCATCGTGGTGCTTTCGGACGCCTTCCTGAGCCGGGCCTGGTGCAGCCAC 946

Qy        721 AGCTTCCGGGAGGGCCTGTGCCGGCTGCTGGAGCTCACCCGCAGACCCATCTTCATCACC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        947 AGCTTCCGGGAGGGCCTGTGCCGGCTGCTGGAGCTCACCCGCAGACCCATCTTCATCACC 1006

Qy        781 TTCGAGGGCCAGAGGCGCGACCCCGCGCACCCGGCGCTCCGCCTGCTGCGCCAGCACCGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1007 TTCGAGGGCCAGAGGCGCGACCCCGCGCACCCGGCGCTCCGCCTGCTGCGCCAGCACCGC 1066

Qy        841 CACCTGGTGACCTTGCTGCTCTGGAGGCCCGGCTCCGTGACTCCTTCCTCCGATTTTTGG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1067 CACCTGGTGACCTTGCTGCTCTGGAGGCCCGGCTCCGTGACTCCTTCCTCCGATTTTTGG 1126

Qy        901 AAAGAAGTGCAGCTGGCGCTGCCGCGGAAGGTGCGGTACAGGCCGGTGGAAGGAGACCCC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1127 AAAGAAGTGCAGCTGGCGCTGCCGCGGAAGGTGCGGTACAGGCCGGTGGAAGGAGACCCC 1186

Qy        961 CAGACGCAGCTGCAGGACGACAAGGACCCCATGCTGATTCTTCGAGGCCGAGTCCCTGAG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1187 CAGACGCAGCTGCAGGACGACAAGGACCCCATGCTGATTCTTCGAGGCCGAGTCCCTGAG 1246

Qy       1021 GGCCGGGCCCTGGACTCAGAGGTGGACCCGGACCCTGAGGGCGACCTGGGTGTCCGGGGG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1247 GGCCGGGCCCTGGACTCAGAGGTGGACCCGGACCCTGAGGGCGACCTGGGTGTCCGGGGG 1306

Qy       1081 CCTGTTTTTGGAGAGCCATCAGCTCCACCGCACACCAGTGGGGTCTCGCTGGGAGAGAGC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1307 CCTGTTTTTGGAGAGCCATCAGCTCCACCGCACACCAGTGGGGTCTCGCTGGGAGAGAGC 1366

Qy       1141 CGGAGCAGCGAAGTGGACGTCTCGGATCTCGGCTCGCGAAACTACAGTGCCCGCACAGAC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1367 CGGAGCAGCGAAGTGGACGTCTCGGATCTCGGCTCGCGAAACTACAGTGCCCGCACAGAC 1426

Qy       1201 TTCTACTGCCTGGTGTCCAAGGATGATATGTAG 1233
              ||||||||||||||||||||||||||||||||
Db       1427 TTCTACTGCCTGGTGTCCAAGGATGATATGTAG 1459


RESULT 4
ADO19086
ID   ADO19086 standard; cDNA; 1695 BP.
XX
AC   ADO19086;
XX
DT   12-AUG-2004  (first entry)
XX
DE   Human PRO polynucleotide #12.
XX
KW   Human; PRO; gene; ss; immune related disorder;
KW   systemic lupus erythematosus; rheumatoid arthritis; osteoarthritis;
KW   juvenile chronic arthritis; systemic sclerosis; Sjogren's syndrome;
KW   vasculitis; sarcoidosis; autoimmune haemolytic anaemia;
KW   autoimmune thrombocytopenia; thyroiditis; diabetes mellitus;
```

KW renal disease; demyelinating disease; central nervous system;
KW peripheral nervous system; demyelinating polyneuropathy;
KW Guillain-Barre syndrome;
KW chronic inflammatory demyelinating polyneuropathy.
XX
OS Homo sapiens.
XX
PN WO2004043361-A2.
XX
PD 27-MAY-2004.
XX
PF 06-NOV-2003; 2003WO-US035268.
XX
PR 08-NOV-2002; 2002US-0425235P.
XX
PA (GETH) GENENTECH INC.
XX
PI Fong S, Dennis K, Clark H, Chiu H, Schoenfeld J, Williams PM;
PI Wood WI, Wu TD;
XX
DR WPI; 2004-420067/39.
DR P-PSDB; ADO19087.
XX
PT Novel PRO polypeptide e.g., PRO69614, PRO71106, or PRO86388 useful for
PT treating an immune related disorder such as systemic lupus erythematosus,
PT rheumatoid arthritis, osteoarthritis, juvenile chronic arthritis or
PT spondyloarthropathy.
XX
PS Claim 1; SEQ ID NO 23; 1731pp; English.
XX
CC The invention relates to human PRO polypeptides and the polynucleotides
CC encoding them. The polypeptides and polynucleotides are useful for
CC treating and diagnosing immune related disorders in mammals. The immune
CC related disorders include systemic lupus erythematosus, rheumatoid
CC arthritis, osteoarthritis, juvenile chronic arthritis, systemic
CC sclerosis, Sjogren's syndrome, vasculitis, sarcoidosis, autoimmune
CC haemolytic anaemia, autoimmune thrombocytopenia, thyroiditis, diabetes
CC mellitus, immune-mediated renal disease, demyelinating diseases of the
CC central or peripheral nervous system, demyelinating polyneuropathy,
CC Guillain-Barre syndrome and chronic inflammatory demyelinating
CC polyneuropathy. This sequence represents a human PRO polynucleotide of
CC the invention.
XX
SQ Sequence 1695 BP; 329 A; 556 C; 510 G; 300 T; 0 U; 0 Other;

Query Match 100.0%; Score 1233; DB 12; Length 1695;
Best Local Similarity 100.0%; Pred. No. 1.1e-236;
Matches 1233; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATGCCAGGTGTCTGTGATAGGGCCCCTGACTTCCTCTCCCCGTCTGAAGACCAGGTGCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 ATGCCAGGTGTCTGTGATAGGGCCCCTGACTTCCTCTCCCCGTCTGAAGACCAGGTGCTG 286

Qy         61 AGGCCTGCCTTGGGCAGCTCAGTGGCTCTGAACTGCACGGCTTGGGTAGTCTCTGGGCCC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 AGGCCTGCCTTGGGCAGCTCAGTGGCTCTGAACTGCACGGCTTGGGTAGTCTCTGGGCCC 346

Qy        121 CACTGCTCCCTGCCTTCAGTCCAGTGGCTGAAAGACGGGCTTCCATTGGGAATTGGGGGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        347 CACTGCTCCCTGCCTTCAGTCCAGTGGCTGAAAGACGGGCTTCCATTGGGAATTGGGGGC 406

Qy        181 CACTACAGCCTCCACGAGTACTCCTGGGTCAAGGCCAACCTGTCAGAGGTGCTTGTGTCC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        407 CACTACAGCCTCCACGAGTACTCCTGGGTCAAGGCCAACCTGTCAGAGGTGCTTGTGTCC 466

Qy        241 AGTGTCCTGGGGGTCAACGTGACCAGCACTGAAGTCTATGGGGCCTTCACCTGCTCCATC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        467 AGTGTCCTGGGGGTCAACGTGACCAGCACTGAAGTCTATGGGGCCTTCACCTGCTCCATC 526

Qy        301 CAGAACATCAGCTTCTCCTCCTTCACTCTTCAGAGAGCTGGCCCTACAAGCCACGTGGCT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        527 CAGAACATCAGCTTCTCCTCCTTCACTCTTCAGAGAGCTGGCCCTACAAGCCACGTGGCT 586

Qy        361 GCGGTGCTGGCCTCCCTCCTGGTCCTGCTGGCCCTGCTGCTGGCCGCCCTGCTCTATGTC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        587 GCGGTGCTGGCCTCCCTCCTGGTCCTGCTGGCCCTGCTGCTGGCCGCCCTGCTCTATGTC 646

Qy        421 AAGTGCCGTCTCAACGTGCTGCTCTGGTACCAGGACGCGTATGGGGAGGTGGAGATAAAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        647 AAGTGCCGTCTCAACGTGCTGCTCTGGTACCAGGACGCGTATGGGGAGGTGGAGATAAAC 706

Qy        481 GACGGGAAGCTCTACGACGCCTACGTCTCCTACAGCGACTGCCCCGAGGACCGCAAGTTC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        707 GACGGGAAGCTCTACGACGCCTACGTCTCCTACAGCGACTGCCCCGAGGACCGCAAGTTC 766

Qy        541 GTGAACTTCATCCTAAAGCCGCAGCTGGAGCGGCGTCGGGGCTACAAGCTCTTCCTGGAC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        767 GTGAACTTCATCCTAAAGCCGCAGCTGGAGCGGCGTCGGGGCTACAAGCTCTTCCTGGAC 826

Qy        601 GACCGCGACCTCCTGCCGCGCGCGCTGAGCCCTCCGCCGACCTCTTGGTGAACCTGAGCCGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        827 GACCGCGACCTCCTGCCGCGCGCGCTGAGCCCTCCGCCGACCTCTTGGTGAACCTGAGCCGC 886

Qy        661 TGCCGACGCCTCATCGTGGTGCTTTCGGACGCCTTCCTGAGCCGGGCCTGGTGCAGCCAC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        887 TGCCGACGCCTCATCGTGGTGCTTTCGGACGCCTTCCTGAGCCGGGCCTGGTGCAGCCAC 946

Qy        721 AGCTTCCGGGAGGGCCTGTGCCGGCTGCTGGAGCTCACCCGCAGACCCATCTTCATCACC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        947 AGCTTCCGGGAGGGCCTGTGCCGGCTGCTGGAGCTCACCCGCAGACCCATCTTCATCACC 1006

Qy        781 TTCGAGGGCCAGAGGCGCGACCCCGCGCACCCGGCGCTCCGCCTGCTGCGCCAGCACCGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1007 TTCGAGGGCCAGAGGCGCGACCCCGCGCACCCGGCGCTCCGCCTGCTGCGCCAGCACCGC 1066

Qy        841 CACCTGGTGACCTTGCTGCTCTGGAGGCCCGGCTCCGTGACTCCTTCCTCCGATTTTTGG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1067 CACCTGGTGACCTTGCTGCTCTGGAGGCCCGGCTCCGTGACTCCTTCCTCCGATTTTTGG 1126

Qy        901 AAAGAAGTGCAGCTGGCGCTGCCGCGGAAGGTGCGGTACAGGCCGGTGGAAGGAGACCCC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1127 AAAGAAGTGCAGCTGGCGCTGCCGCGGAAGGTGCGGTACAGGCCGGTGGAAGGAGACCCC 1186

Qy        961 CAGACGCAGCTGCAGGACGACAAGGACCCCATGCTGATTCTTCGAGGCCGAGTCCCTGAG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1187 CAGACGCAGCTGCAGGACGACAAGGACCCCATGCTGATTCTTCGAGGCCGAGTCCCTGAG 1246

Qy       1021 GGCCGGGCCCTGGACTCAGAGGTGGACCCGGACCCTGAGGGCGACCTGGGTGTCCGGGGG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1247 GGCCGGGCCCTGGACTCAGAGGTGGACCCGGACCCTGAGGGCGACCTGGGTGTCCGGGGG 1306

Qy       1081 CCTGTTTTTGGAGAGCCATCAGCTCCACCGCACACCAGTGGGGTCTCGCTGGGAGAGAGC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1307 CCTGTTTTTGGAGAGCCATCAGCTCCACCGCACACCAGTGGGGTCTCGCTGGGAGAGAGC 1366

Qy       1141 CGGAGCAGCGAAGTGGACGTCTCGGATCTCGGCTCGCGAAACTACAGTGCCCGCACAGAC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1367 CGGAGCAGCGAAGTGGACGTCTCGGATCTCGGCTCGCGAAACTACAGTGCCCGCACAGAC 1426

Qy       1201 TTCTACTGCCTGGTGTCCAAGGATGATATGTAG 1233
              ||||||||||||||||||||||||||||||||
Db       1427 TTCTACTGCCTGGTGTCCAAGGATGATATGTAG 1459
```

RESULT 5
ADP22969
ID ADP22969 standard; cDNA; 1695 BP.
XX
AC ADP22969;
XX
DT 18-NOV-2004 (first entry)
XX
DE PRO polypeptide encoding cDNA SEQ ID NO:63.
XX
KW ss; gene; PRO; antiinflammatory; antiarthritic; antirheumatic;
KW immunosuppressive; osteopathic; antidiabetic; dermatological;

KW antipsoriatic; antiallergic; antiasthmatic; hepatotropic; respiratory;
KW gene therapy; immune system.
XX
OS Unidentified.
XX
PN WO2004041170-A2.
XX
PD 21-MAY-2004.
XX
PF 30-OCT-2003; 2003WO-US034312.
XX
PR 01-NOV-2002; 2002US-0423394P.
XX
PA (GETH) GENENTECH INC.
XX
PI Clark H, Schoenfeld J, Van Lookeren M, Williams PM, Wood WI;
PI Wu TD;
XX
DR WPI; 2004-419628/39.
DR P-PSDB; ADP22970.
XX
PT New PRO polypeptides and polynucleotides, useful for treating e.g.
PT erythematosus, rheumatoid arthritis, diabetes mellitus, immune-mediated
PT renal disease, or demyelinating diseases of the central or peripheral
PT nervous system.
XX
PS Claim 1; SEQ ID NO 63; 2940pp; English.
XX
CC The invention relates to a novel isolated nucleic acid and the PRO
CC polypeptide encoded by it. A protein of the invention has
CC antiinflammatory, antiarthritic, antirheumatic, immunosuppressive,
CC osteopathic, antidiabetic, dermatological, antipsoriatic, antiallergic,
CC antiasthmatic, hepatotropic, and respiratory activity. A polynucleotide
CC of the invention may have a use in gene therapy. The PRO polypeptide, its
CC agonist, antagonist, or antibody that specifically binds to the
CC polypeptide is useful for treating an immune related disorder such as
CC systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis,
CC juvenile chronic arthritis, a spondyloarthropathy, systemic sclerosis, an
CC idiopathic inflammatory myopathy, Sjogren's syndrome, systemic
CC vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune
CC thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal
CC disease, a demyelinating disease of the central or peripheral nervous
CC system, idiopathic demyelinating polyneuropathy, Guillain-Barre syndrome,
CC a chronic inflammatory demyelinating polyneuropathy, a hepatobiliary
CC disease, infectious or autoimmune chronic active hepatitis, primary
CC biliary cirrhosis, granulomatous hepatitis, sclerosing cholangitis,
CC inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's
CC disease, an autoimmune or immune-mediated skin disease, a bullous skin
CC disease, erythema multiforme, contact dermatitis, psoriasis, an allergic
CC disease, asthma, allergic rhinitis, atopic dermatitis, food
CC hypersensitivity, urticaria, an immunologic disease of the lung,
CC eosinophilic pneumonia, idiopathic pulmonary fibrosis, hypersensitivity
CC pneumonitis, a transplantation associated disease, graft rejection or
CC graft-versus-host disease. The present sequence encodes a PRO protein of
CC the invention.
XX
SQ Sequence 1695 BP; 329 A; 556 C; 510 G; 300 T; 0 U; 0 Other;

Query Match 100.0%; Score 1233; DB 13; Length 1695;
Best Local Similarity 100.0%; Pred. No. 1.1e-236;
Matches 1233; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATGCCAGGTGTCTGTGATAGGGCCCCTGACTTCCTCTCCCCGTCTGAAGACCAGGTGCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        227 ATGCCAGGTGTCTGTGATAGGGCCCCTGACTTCCTCTCCCCGTCTGAAGACCAGGTGCTG 286

Qy         61 AGGCCTGCCTTGGGCAGCTCAGTGGCTCTGAACTGCACGGCTTGGGTAGTCTCTGGGCCC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        287 AGGCCTGCCTTGGGCAGCTCAGTGGCTCTGAACTGCACGGCTTGGGTAGTCTCTGGGCCC 346

Qy        121 CACTGCTCCCTGCCTTCAGTCCAGTGGCTGAAAGACGGGCTTCCATTGGGAATTGGGGGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        347 CACTGCTCCCTGCCTTCAGTCCAGTGGCTGAAAGACGGGCTTCCATTGGGAATTGGGGGC 406

Qy        181 CACTACAGCCTCCACGAGTACTCCTGGGTCAAGGCCAACCTGTCAGAGGTGCTTGTGTCC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        407 CACTACAGCCTCCACGAGTACTCCTGGGTCAAGGCCAACCTGTCAGAGGTGCTTGTGTCC 466

Qy        241 AGTGTCCTGGGGGTCAACGTGACCAGCACTGAAGTCTATGGGGCCTTCACCTGCTCCATC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        467 AGTGTCCTGGGGGTCAACGTGACCAGCACTGAAGTCTATGGGGCCTTCACCTGCTCCATC 526

Qy        301 CAGAACATCAGCTTCTCCTCCTTCACTCTTCAGAGAGCTGGCCCTACAAGCCACGTGGCT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        527 CAGAACATCAGCTTCTCCTCCTTCACTCTTCAGAGAGCTGGCCCTACAAGCCACGTGGCT 586

Qy        361 GCGGTGCTGGCCTCCCTCCTGGTCCTGCTGGCCCTGCTGCTGGCCGCCCTGCTCTATGTC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        587 GCGGTGCTGGCCTCCCTCCTGGTCCTGCTGGCCCTGCTGCTGGCCGCCCTGCTCTATGTC 646

Qy        421 AAGTGCCGTCTCAACGTGCTGCTCTGGTACCAGGACGCGTATGGGGAGGTGGAGATAAAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        647 AAGTGCCGTCTCAACGTGCTGCTCTGGTACCAGGACGCGTATGGGGAGGTGGAGATAAAC 706

Qy        481 GACGGGAAGCTCTACGACGCCTACGTCTCCTACAGCGACTGCCCCGAGGACCGCAAGTTC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        707 GACGGGAAGCTCTACGACGCCTACGTCTCCTACAGCGACTGCCCCGAGGACCGCAAGTTC 766

Qy        541 GTGAACTTCATCCTAAAGCCGCAGCTGGAGCGGCGTCGGGGCTACAAGCTCTTCCTGGAC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        767 GTGAACTTCATCCTAAAGCCGCAGCTGGAGCGGCGTCGGGGCTACAAGCTCTTCCTGGAC 826

Qy        601 GACCGCGACCTCCTGCCGCGCGCTGAGCCCTCCGCCGACCTCTTGGTGAACCTGAGCCGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        827 GACCGCGACCTCCTGCCGCGCGCTGAGCCCTCCGCCGACCTCTTGGTGAACCTGAGCCGC 886

Qy        661 TGCCGACGCCTCATCGTGGTGCTTTCGGACGCCTTCCTGAGCCGGGCCTGGTGCAGCCAC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        887 TGCCGACGCCTCATCGTGGTGCTTTCGGACGCCTTCCTGAGCCGGGCCTGGTGCAGCCAC 946

Qy        721 AGCTTCCGGGAGGGCCTGTGCCGGCTGCTGGAGCTCACCCGCAGACCCATCTTCATCACC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        947 AGCTTCCGGGAGGGCCTGTGCCGGCTGCTGGAGCTCACCCGCAGACCCATCTTCATCACC 1006

Qy        781 TTCGAGGGCCAGAGGCGCGACCCCGCGCACCCGGCGCTCCGCCTGCTGCGCCAGCACCGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1007 TTCGAGGGCCAGAGGCGCGACCCCGCGCACCCGGCGCTCCGCCTGCTGCGCCAGCACCGC 1066

Qy        841 CACCTGGTGACCTTGCTGCTCTGGAGGCCCGGCTCCGTGACTCCTTCCTCCGATTTTTGG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1067 CACCTGGTGACCTTGCTGCTCTGGAGGCCCGGCTCCGTGACTCCTTCCTCCGATTTTTGG 1126

Qy        901 AAAGAAGTGCAGCTGGCGCTGCCGCGGAAGGTGCGGTACAGGCCGGTGGAAGGAGACCCC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1127 AAAGAAGTGCAGCTGGCGCTGCCGCGGAAGGTGCGGTACAGGCCGGTGGAAGGAGACCCC 1186

Qy        961 CAGACGCAGCTGCAGGACGACAAGGACCCCATGCTGATTCTTCGAGGCCGAGTCCCTGAG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1187 CAGACGCAGCTGCAGGACGACAAGGACCCCATGCTGATTCTTCGAGGCCGAGTCCCTGAG 1246

Qy       1021 GGCCGGGCCCTGGACTCAGAGGTGGACCCGGACCCTGAGGGCGACCTGGGTGTCCGGGGG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1247 GGCCGGGCCCTGGACTCAGAGGTGGACCCGGACCCTGAGGGCGACCTGGGTGTCCGGGGG 1306

Qy       1081 CCTGTTTTTGGAGAGCCATCAGCTCCACCGCACACCAGTGGGGTCTCGCTGGGAGAGAGC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1307 CCTGTTTTTGGAGAGCCATCAGCTCCACCGCACACCAGTGGGGTCTCGCTGGGAGAGAGC 1366

Qy       1141 CGGAGCAGCGAAGTGGACGTCTCGGATCTCGGCTCGCGAAACTACAGTGCCCGCACAGAC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1367 CGGAGCAGCGAAGTGGACGTCTCGGATCTCGGCTCGCGAAACTACAGTGCCCGCACAGAC 1426

Qy       1201 TTCTACTGCCTGGTGTCCAAGGATGATATGTAG 1233
              |||||||||||||||||||||||||||||||||
Db       1427 TTCTACTGCCTGGTGTCCAAGGATGATATGTAG 1459
```

```
RESULT 6
AAF27859
ID   AAF27859 standard; cDNA; 1659 BP.
XX
AC   AAF27859;
XX
DT   30-MAR-2001  (first entry)
XX
DE   Human NOV9 cDNA.
XX
KW   Human; NOVX; antiinflammatory; cytostatic; neuroprotective;
KW   cerebroprotective; immunomodulator; vulnerary; vasotropic; gene therapy;
KW   hyperplasia; tumour; restenosis; psoriasis; Dupuytren's contracture;
KW   diabetes; rheumatoid arthritis; cerebral oedema; Alzheimer's disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200075321-A2.
XX
PD   14-DEC-2000.
XX
PF   01-JUN-2000; 2000WO-US015303.
XX
PR   03-JUN-1999;   99US-0137322P.
PR   16-MAR-2000; 2000US-0189810P.
PR   22-MAR-2000; 2000US-0191158P.
PR   30-MAR-2000; 2000US-0193086P.
PR   31-MAY-2000; 2000US-00137322.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Fernandes E,  Herrman J,  Vernet C;
XX
DR   WPI; 2001-102403/11.
DR   P-PSDB; AAB61137.
XX
PT   New NOVX polypeptides and polynucleotides, useful in gene therapy, as a
PT   diagnostic marker, protein therapeutic, antibody or small molecule drug
PT   target for treating immune, proliferative and metabolic diseases and
PT   wound healing.
XX
PS   Claim 8; Page 36-38; 194pp; English.
XX
CC   The present sequence encodes a new isolated polypeptide (NOVX). The NOVX
CC   polypeptides, NOVX nucleic acids, and anti-NOVX antibodies are useful for
CC   treating or preventing NOVX-associated disorders. They are also useful
CC   for determining the presence of or a predisposition to a disease
CC   associated with altered levels of the NOVX polypeptide or nucleic acid.
CC   These NOVX-associated disorders include hyperplasias, tumours,
CC   restenosis, psoriasis, Dupuytren's contracture, diabetic complications,
CC   rheumatoid arthritis, cerebral lesions, diabetic neuropathies, cerebral
CC   oedema, senile dementia or Alzheimer's disease. The NOVX polynucleotides
CC   are especially useful in gene therapy. Specifically, NOVX is useful as a
CC   diagnostic marker or prognostic marker, protein therapeutic and antibody
CC   target or small molecule drug target to treat disorders in the immune
CC   response pathway, thyroid and metabolic diseases, bone metabolic
CC   disorders, diseases of the pancreas (e.g. diabetes or digestive
CC   disorders), proliferative diseases, or tissue regeneration and
CC   development (e.g. wound healing or treatment of burns)
XX
SQ   Sequence 1659 BP; 296 A; 551 C; 504 G; 308 T; 0 U; 0 Other;

  Query Match             99.9%;  Score 1231.4;  DB 4;  Length 1659;
  Best Local Similarity   99.9%;  Pred. No. 2.3e-236;
  Matches 1232;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 ATGCCAGGTGTCTGTGATAGGGCCCCTGACTTCCTCTCCCCGTCTGAAGACCAGGTGCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 ATGCCAGGTGTCTGTGATAGGGCCCCTGACTTCCTCTCCCCGTCTGAAGACCAGGTGCTG 303

Qy         61 AGGCCTGCCTTGGGCAGCTCAGTGGCTCTGAACTGCACGGCTTGGGTAGTCTCTGGGCCC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        304 AGGCCTGCCTTGGGCAGCTCAGTGGCTCTGAACTGCACGGCTTGGGTAGTCTCTGGGCCC 363

Qy        121 CACTGCTCCCTGCCTTCAGTCCAGTGGCTGAAAGACGGGCTTCCATTGGGAATTGGGGGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        364 CACTGCTCCCTGCCTTCAGTCCAGTGGCTGAAAGACGGGCTTCCATTGGGAATTGGGGGC 423

Qy        181 CACTACAGCCTCCACGAGTACTCCTGGGTCAAGGCCAACCTGTCAGAGGTGCTTGTGTCC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        424 CACTACAGCCTCCACGAGTACTCCTGGGTCAAGGCCAACCTGTCAGAGGTGCTTGTGTCC 483

Qy        241 AGTGTCCTGGGGGTCAACGTGACCAGCACTGAAGTCTATGGGGCCTTCACCTGCTCCATC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        484 AGTGTCCTGGGGGTCAACGTGACCAGCACTGAAGTCTATGGGGCCTTCACCTGCTCCATC 543

Qy        301 CAGAACATCAGCTTCTCCTCCTTCACTCTTCAGAGAGCTGGCCCTACAAGCCACGTGGCT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        544 CAGAACATCAGCTTCTCCTCCTTCACTCTTCAGAGAGCTGGCCCTACAAGCCACGTGGCT 603

Qy        361 GCGGTGCTGGCCTCCCTCCTGGTCCTGCTGGCCCTGCTGCTGGCCGCCCTGCTCTATGTC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        604 GCGGTGCTGGCCTCCCTCCTGGTCCTGCTGGCCCTGCTGCTGGCCGCCCTGCTCTATGTC 663

Qy        421 AAGTGCCGTCTCAACGTGCTGCTCTGGTACCAGGACGCGTATGGGGAGGTGGAGATAAAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        664 AAGTGCCGTCTCAACGTGCTGCTCTGGTACCAGGACGCGTATGGGGAGGTGGAGATAAAC 723

Qy        481 GACGGGAAGCTCTACGACGCCTACGTCTCCTACAGCGACTGCCCCGAGGACCGCAAGTTC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        724 GACGGGAAGCTCTACGACGCCTACGTCTCCTACAGCGACTGCCCCGAGGACCGCAAGTTC 783

Qy        541 GTGAACTTCATCCTAAAGCCGCAGCTGGAGCGGCGTCGGGGCTACAAGCTCTTCCTGGAC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        784 GTGAACTTCATCCTAAAGCCGCAGCTGGAGCGGCGTCGGGGCTACAAGCTCTTCCTGGAC 843

Qy        601 GACCGCGACCTCCTGCCGCGCGCTGAGCCCTCCGCCGACCTCTTGGTGAACCTGAGCCGC 660
              ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        844 GACCGCGACCTCCTGCCCCGCGCTGAGCCCTCCGCCGACCTCTTGGTGAACCTGAGCCGC 903

Qy        661 TGCCGACGCCTCATCGTGGTGCTTTCGGACGCCTTCCTGAGCCGGGCCTGGTGCAGCCAC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        904 TGCCGACGCCTCATCGTGGTGCTTTCGGACGCCTTCCTGAGCCGGGCCTGGTGCAGCCAC 963

Qy        721 AGCTTCCGGGAGGGCCTGTGCCGGCTGCTGGAGCTCACCCGCAGACCCATCTTCATCACC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        964 AGCTTCCGGGAGGGCCTGTGCCGGCTGCTGGAGCTCACCCGCAGACCCATCTTCATCACC 1023

Qy        781 TTCGAGGGCCAGAGGCGCGACCCCGCGCACCCGGCGCTCCGCCTGCTGCGCCAGCACCGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1024 TTCGAGGGCCAGAGGCGCGACCCCGCGCACCCGGCGCTCCGCCTGCTGCGCCAGCACCGC 1083

Qy        841 CACCTGGTGACCTTGCTGCTCTGGAGGCCCGGCTCCGTGACTCCTTCCTCCGATTTTTGG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1084 CACCTGGTGACCTTGCTGCTCTGGAGGCCCGGCTCCGTGACTCCTTCCTCCGATTTTTGG 1143

Qy        901 AAAGAAGTGCAGCTGGCGCTGCCGCGGAAGGTGCGGTACAGGCCGGTGGAAGGAGACCCC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1144 AAAGAAGTGCAGCTGGCGCTGCCGCGGAAGGTGCGGTACAGGCCGGTGGAAGGAGACCCC 1203

Qy        961 CAGACGCAGCTGCAGGACGACAAGGACCCCATGCTGATTCTTCGAGGCCGAGTCCCTGAG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1204 CAGACGCAGCTGCAGGACGACAAGGACCCCATGCTGATTCTTCGAGGCCGAGTCCCTGAG 1263

Qy       1021 GGCCGGGCCCTGGACTCAGAGGTGGACCCGGACCCTGAGGGCGACCTGGGTGTCCGGGGG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1264 GGCCGGGCCCTGGACTCAGAGGTGGACCCGGACCCTGAGGGCGACCTGGGTGTCCGGGGG 1323

Qy       1081 CCTGTTTTTGGAGAGCCATCAGCTCCACCGCACACCAGTGGGGTCTCGCTGGGAGAGAGC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1324 CCTGTTTTTGGAGAGCCATCAGCTCCACCGCACACCAGTGGGGTCTCGCTGGGAGAGAGC 1383

Qy       1141 CGGAGCAGCGAAGTGGACGTCTCGGATCTCGGCTCGCGAAACTACAGTGCCCGCACAGAC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1384 CGGAGCAGCGAAGTGGACGTCTCGGATCTCGGCTCGCGAAACTACAGTGCCCGCACAGAC 1443
```

```
Qy       1201 TTCTACTGCCTGGTGTCCAAGGATGATATGTAG 1233
             |||||||||||||||||||||||||||||||||
Db       1444 TTCTACTGCCTGGTGTCCAAGGATGATATGTAG 1476


RESULT 7
ACD40262
ID   ACD40262 standard; DNA; 1659 BP.
XX
AC   ACD40262;
XX
DT   03-SEP-2003  (first entry)
XX
DE   Human breast tumour associated protein 47-like polypeptide NOV9 DNA.
XX
KW   Tissue typing; cancer; breast cancer; colon cancer; lung cancer; sarcoma;
KW   pancreatic cancer; uterine cancer; organ transplantation disorder; ds;
KW   cardiovascular disease; melanoma; atherosclerosis; diabetes mellitus;
KW   ischaemic heart disease; haemorrhage; peripheral vascular disease; gene;
KW   thrombosis; hypertension; systemic lupus erythematosus; haematopoiesis;
KW   tissue regeneration; wound healing; hyperproliferative disorder;
KW   psoriasis; neural disorder; Parkinson's disease; Alzheimer's disease;
KW   Huntington's disease; multiple sclerosis; amyotrophic lateral sclerosis;
KW   ALS; peripheral neuropathy; nervous system tumour; neurotoxin; tremor;
KW   neuropathy; acute brain injury; peripheral nerve trauma; human; NOVX;
KW   gene therapy; epilepsy; breast tumour associated protein 47.
XX
OS   Homo sapiens.
XX
PN   US2003027158-A1.
XX
PD   06-FEB-2003.
XX
PF   15-OCT-2001; 2001US-00977418.
XX
PR   03-JUN-1999;   99US-0137322P.
PR   16-MAR-2000; 2000US-0189810P.
PR   22-MAR-2000; 2000US-0191158P.
PR   30-MAR-2000; 2000US-0193086P.
PR   03-MAY-2000; 2000US-0201388P.
PR   31-MAY-2000; 2000US-00584411.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Fernandes E,  Herrman J,  Vernet C;
XX
DR   WPI; 2003-492028/46.
DR   P-PSDB; ABO23241.
XX
PT   New nucleic acid sequence encoding a human breast tumor-associated
PT   protein 47-like polypeptide, useful for treating cardiovascular
PT   disorders, neural disorders, diabetes mellitus and cancers.
XX
PS   Disclosure; Page 23-25; 100pp; English.
XX
CC   The invention relates to a new isolated NOV4 nucleic acid. The nucleic
CC   acid is useful for identifying a compound that binds the nucleic acid.
CC   The nucleic acid is useful in gene therapy, in screening assays, in
CC   detection assays e.g. chromosomal mapping, cell and tissue typing and
CC   forensic biology, predictive medicine e.g. diagnostic assays, prognostic
CC   assays, monitoring clinical trials and pharmacogenomics and methods of
CC   treatment including therapeutic and prophylactic. The nucleic acid is
CC   also useful for expressing NOVX protein. The nucleic acid is also useful
CC   to provide polynucleotide reagents e.g. labelled probes that are useful
CC   in an in situ hybridisation technique, for identifying a specific tissue
CC   (for example brain tissue) and for use in forensic science. The nucleic
CC   acid is also useful for mapping genes on a chromosome and thus locating
CC   gene regions associated with genetic disease, identifying an individual
CC   from a minute biological sample and to aid in forensic identification of
CC   biological sample. The nucleic acid is also useful for treating cancer,
CC   especially cancers of the breast, colon, lung, pancreas or uterus, or a
CC   melanoma or sarcoma. The nucleic acid is also useful for treating
CC   disorders related to organ transplantation, cardiovascular diseases,
CC   atherosclerosis, ischaemic heart disease, haemorrhage, diabetes mellitus,
CC   peripheral vascular disease, thrombosis, hypertension and systemic lupus
CC   erythematosus. NOVX protein encoded by the nucleic acid is useful for
CC   regulating haematopoiesis, for regeneration of bone, cartilage, tendon
CC   ligament and/or nerve tissue growth or regeneration and for wound
CC   healing. The nucleic acid is also useful for treating infections,
CC   hyperproliferative disorders e.g. psoriasis, and neural disorders
CC   including Parkinson's disease, Alzheimer's disease, Huntington's disease,
CC   multiple sclerosis, amyotrophic lateral sclerosis (ALS), peripheral
CC   neuropathy, tumours of the nervous system, exposure to neurotoxins, acute
CC   brain injury, peripheral nerve trauma or injury and other neuropathies,
CC   epilepsy, and/or tremors. The present sequence represents DNA encoding a
CC   human breast tumour associated protein 47-like polypeptide
XX
SQ   Sequence 1659 BP; 295 A; 552 C; 504 G; 308 T; 0 U; 0 Other;

  Query Match             99.9%;  Score 1231.4;  DB 9;  Length 1659;
  Best Local Similarity   99.9%;  Pred. No. 2.3e-236;
  Matches 1232;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 ATGCCAGGTGTCTGTGATAGGGCCCCTGACTTCCTCTCCCCGTCTGAAGACCAGGTGCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 ATGCCAGGTGTCTGTGATAGGGCCCCTGACTTCCTCTCCCCGTCTGAAGACCAGGTGCTG 303

Qy         61 AGGCCTGCCTTGGGCAGCTCAGTGGCTCTGAACTGCACGGCTTGGGTAGTCTCTGGGCCC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        304 AGGCCTGCCTTGGGCAGCTCAGTGGCTCTGAACTGCACGGCTTGGGTAGTCTCTGGGCCC 363

Qy        121 CACTGCTCCCTGCCTTCAGTCCAGTGGCTGAAAGACGGGCTTCCATTGGGAATTGGGGGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        364 CACTGCTCCCTGCCTTCAGTCCAGTGGCTGAAAGACGGGCTTCCATTGGGAATTGGGGGC 423

Qy        181 CACTACAGCCTCCACGAGTACTCCTGGGTCAAGGCCAACCTGTCAGAGGTGCTTGTGTCC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        424 CACTACAGCCTCCACGAGTACTCCTGGGTCAAGGCCAACCTGTCAGAGGTGCTTGTGTCC 483

Qy        241 AGTGTCCTGGGGGTCAACGTGACCAGCACTGAAGTCTATGGGGCCTTCACCTGCTCCATC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        484 AGTGTCCTGGGGGTCAACGTGACCAGCACTGAAGTCTATGGGGCCTTCACCTGCTCCATC 543

Qy        301 CAGAACATCAGCTTCTCCTCCTTCACTCTTCAGAGAGCTGGCCCTACAAGCCACGTGGCT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        544 CAGAACATCAGCTTCTCCTCCTTCACTCTTCAGAGAGCTGGCCCTACAAGCCACGTGGCT 603

Qy        361 GCGGTGCTGGCCTCCCTCCTGGTCCTGCTGGCCCTGCTGCTGGCCGCCCTGCTCTATGTC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        604 GCGGTGCTGGCCTCCCTCCTGGTCCTGCTGGCCCTGCTGCTGGCCGCCCTGCTCTATGTC 663

Qy        421 AAGTGCCGTCTCAACGTGCTGCTCTGGTACCAGGACGCGTATGGGGAGGTGGAGATAAAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        664 AAGTGCCGTCTCAACGTGCTGCTCTGGTACCAGGACGCGTATGGGGAGGTGGAGATAAAC 723

Qy        481 GACGGGAAGCTCTACGACGCCTACGTCTCCTACAGCGACTGCCCCGAGGACCGCAAGTTC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        724 GACGGGAAGCTCTACGACGCCTACGTCTCCTACAGCGACTGCCCCGAGGACCGCAAGTTC 783

Qy        541 GTGAACTTCATCCTAAAGCCGCAGCTGGAGCGGCGTCGGGGCTACAAGCTCTTCCTGGAC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        784 GTGAACTTCATCCTAAAGCCGCAGCTGGAGCGGCGTCGGGGCTACAAGCTCTTCCTGGAC 843

Qy        601 GACCGCGACCTCCTGCCGCGCGCTGAGCCCTCCGCCGACCTCTTGGTGAACCTGAGCCGC 660
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        844 GACCGCGACCTCCTGCCCCGCGCTGAGCCCTCCGCCGACCTCTTGGTGAACCTGAGCCGC 903

Qy        661 TGCCGACGCCTCATCGTGGTGCTTTCGGACGCCTTCCTGAGCCGGGCCTGGTGCAGCCAC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        904 TGCCGACGCCTCATCGTGGTGCTTTCGGACGCCTTCCTGAGCCGGGCCTGGTGCAGCCAC 963

Qy        721 AGCTTCCGGGAGGGCCTGTGCCGGCTGCTGGAGCTCACCCGCAGACCCATCTTCATCACC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        964 AGCTTCCGGGAGGGCCTGTGCCGGCTGCTGGAGCTCACCCGCAGACCCATCTTCATCACC 1023
```

```
Qy        781 TTCGAGGGCCAGAGGCGCGACCCCGCGCACCCGGCGCTCCGCCTGCTGCGCCAGCACCGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1024 TTCGAGGGCCAGAGGCGCGACCCCGCGCACCCGGCGCTCCGCCTGCTGCGCCAGCACCGC 1083

Qy        841 CACCTGGTGACCTTGCTGCTCTGGAGGCCCGGCTCCGTGACTCCTTCCTCCGATTTTTGG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1084 CACCTGGTGACCTTGCTGCTCTGGAGGCCCGGCTCCGTGACTCCTTCCTCCGATTTTTGG 1143

Qy        901 AAAGAAGTGCAGCTGGCGCTGCCGCGGAAGGTGCGGTACAGGCCGGTGGAAGGAGACCCC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1144 AAAGAAGTGCAGCTGGCGCTGCCGCGGAAGGTGCGGTACAGGCCGGTGGAAGGAGACCCC 1203

Qy        961 CAGACGCAGCTGCAGGACGACAAGGACCCCATGCTGATTCTTCGAGGCCGAGTCCCTGAG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1204 CAGACGCAGCTGCAGGACGACAAGGACCCCATGCTGATTCTTCGAGGCCGAGTCCCTGAG 1263

Qy       1021 GGCCGGGCCCTGGACTCAGAGGTGGACCCGGACCCTGAGGGCGACCTGGGTGTCCGGGGG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1264 GGCCGGGCCCTGGACTCAGAGGTGGACCCGGACCCTGAGGGCGACCTGGGTGTCCGGGGG 1323

Qy       1081 CCTGTTTTTGGAGAGCCATCAGCTCCACCGCACACCAGTGGGGTCTCGCTGGGAGAGAGC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1324 CCTGTTTTTGGAGAGCCATCAGCTCCACCGCACACCAGTGGGGTCTCGCTGGGAGAGAGC 1383

Qy       1141 CGGAGCAGCGAAGTGGACGTCTCGGATCTCGGCTCGCGAAACTACAGTGCCCGCACAGAC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1384 CGGAGCAGCGAAGTGGACGTCTCGGATCTCGGCTCGCGAAACTACAGTGCCCGCACAGAC 1443

Qy       1201 TTCTACTGCCTGGTGTCCAAGGATGATATGTAG 1233
              ||||||||||||||||||||||||||||||||
Db       1444 TTCTACTGCCTGGTGTCCAAGGATGATATGTAG 1476
```

RESULT 8
ADM56382
ID ADM56382 standard; cDNA; 1659 BP.
XX
AC ADM56382;
XX
DT 03-JUN-2004 (first entry)
XX
DE Human cDNA encoding cell adhesion molecule NOV9.
XX
KW Human; ss; gene; cell adhesion molecule; NOVX; cancer; leukaemia;
KW lymphoma; melanoma; neurological disorder; epilepsy;
KW ischaemic cerebrovascular disease; stroke; Alzheimer's disease;
KW Pick's disease; vesicular transport disorder; cystic fibrosis;
KW diabetes mellitus; Grave's disease; goiter; gastrointestinal disorder;
KW ulcerative colitis; gastric ulcer; duodenal disorder; autoimmune disease;
KW allergic reaction; autoimmune haemolytic anaemia; rheumatoid arthritis;
KW viral infection; bacterial infection; fungal infection;
KW helminthic infection; protozoal infections.
XX
OS Homo sapiens.
XX
PN US2003082554-A1.
XX
PD 01-MAY-2003.
XX
PF 15-OCT-2001; 2001US-00977033.
XX
PR 03-JUN-1999; 99US-0137322P.
PR 16-MAR-2000; 2000US-0189810P.
PR 22-MAR-2000; 2000US-0191158P.
PR 30-MAR-2000; 2000US-0193086P.
PR 03-MAY-2000; 2000US-0201388P.
PR 31-MAY-2000; 2000US-00584411.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Fernandes E, Herrman J, Vernet C;
XX
DR WPI; 2003-616079/58.
DR P-PSDB; ADM56381.
XX
PT New nucleic acids encoding human cell adhesion molecule-like proteins,
PT useful for treating e.g. cancers, neurological disorders, viral,
PT bacterial, fungal, helminthic and protozoal infections.
XX
PS Example 6; SEQ ID NO 17; 78pp; English.
XX
CC The invention relates to an isolated nucleic acid encoding a human cell
CC adhesion molecule-like protein, comprising a sequence encoding a
CC polypeptide having a sequence appearing as ADM56387, a sequence at least
CC 90% identical to the nucleic acid, a sequence encoding a polypeptide
CC having conservative amino acid substitutions to the protein or a fragment
CC comprising at least 20 nucleotides. Also included are an oligonucleotide
CC sequence that is complementary to (and hybridises under stringent
CC conditions with) the nucleic acid (or a portion of it), a vector
CC comprising the nucleic acid, a cell comprising the vector, a
CC pharmaceutical composition comprising the nucleic acid and a
CC pharmaceutical carrier, a process for producing the polypeptide, a
CC process for identifying a compound that binds the nucleic acid, and a
CC compound identified by the process. Disclosed as new are the cDNA and
CC proteins for novel cell adhesion molecules (termed NOVX, being NOV1-23).
CC The NOVX polypeptide, nucleic acid or antibody are useful in the
CC manufacture of a medicament for treating a syndrome associated with a
CC human disease selected from NOVX-associated disorder, such as cancers
CC (e.g. leukaemia, lymphoma, melanoma or cancer of the liver, lung, muscle,
CC ovary, testis and uterus), neurological disorders (e.g. epilepsy,
CC ischaemic cerebrovascular disease, stroke, Alzheimer's disease or Pick's
CC disease), disorders of vesicular transport (e.g. cystic fibrosis,
CC diabetes mellitus, Grave's disease, or goiter), gastrointestinal
CC disorders (e.g. ulcerative colitis, or gastric and duodenal disorders),
CC autoimmune diseases (e.g. allergic reactions, autoimmune haemolytic
CC anaemia, or rheumatoid arthritis), viral, bacterial, fungal, helminthic
CC and protozoal infections. The polypeptides can be used as immunogens to
CC produce antibodies and as vaccines. The sequences may further be used in
CC chromosome mapping, identifying individual from minute biological samples
CC (tissue typing), and in forensic identification of a biological sample.
CC NOTE: The authors have mis-labelled the sequences as they appear on pages
CC 12-28 of the patent, it is clear from table 3, the examples and the
CC claims that the SEQ ID numbers for the cDNAs should be the odd numbers
CC from 1-45 and the proteins should be the even numbers from 2-46. The
CC present sequence encodes a NOVX cell adhesion molecule of the invention.
XX
SQ Sequence 1659 BP; 296 A; 551 C; 504 G; 308 T; 0 U; 0 Other;

Query Match 99.9%; Score 1231.4; DB 11; Length 1659;
Best Local Similarity 99.9%; Pred. No. 2.3e-236;
Matches 1232; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 ATGCCAGGTGTCTGTGATAGGGCCCCTGACTTCCTCTCCCCGTCTGAAGACCAGGTGCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 ATGCCAGGTGTCTGTGATAGGGCCCCTGACTTCCTCTCCCCGTCTGAAGACCAGGTGCTG 303

Qy         61 AGGCCTGCCTTGGGCAGCTCAGTGGCTCTGAACTGCACGGCTTGGGTAGTCTCTGGGCCC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        304 AGGCCTGCCTTGGGCAGCTCAGTGGCTCTGAACTGCACGGCTTGGGTAGTCTCTGGGCCC 363

Qy        121 CACTGCTCCCTGCCTTCAGTCCAGTGGCTGAAAGACGGGCTTCCATTGGGAATTGGGGGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        364 CACTGCTCCCTGCCTTCAGTCCAGTGGCTGAAAGACGGGCTTCCATTGGGAATTGGGGGC 423

Qy        181 CACTACAGCCTCCACGAGTACTCCTGGGTCAAGGCCAACCTGTCAGAGGTGCTTGTGTCC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        424 CACTACAGCCTCCACGAGTACTCCTGGGTCAAGGCCAACCTGTCAGAGGTGCTTGTGTCC 483

Qy        241 AGTGTCCTGGGGGTCAACGTGACCAGCACTGAAGTCTATGGGGCCTTCACCTGCTCCATC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        484 AGTGTCCTGGGGGTCAACGTGACCAGCACTGAAGTCTATGGGGCCTTCACCTGCTCCATC 543

Qy        301 CAGAACATCAGCTTCTCCTCCTTCACTCTTCAGAGAGCTGGCCCTACAAGCCACGTGGCT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        544 CAGAACATCAGCTTCTCCTCCTTCACTCTTCAGAGAGCTGGCCCTACAAGCCACGTGGCT 603
```

```
Qy        361 GCGGTGCTGGCCTCCCTCCTGGTCCTGCTGGCCCTGCTGCTGGCCGCCCTGCTCTATGTC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        604 GCGGTGCTGGCCTCCCTCCTGGTCCTGCTGGCCCTGCTGCTGGCCGCCCTGCTCTATGTC 663

Qy        421 AAGTGCCGTCTCAACGTGCTGCTCTGGTACCAGGACGCGTATGGGGAGGTGGAGATAAAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        664 AAGTGCCGTCTCAACGTGCTGCTCTGGTACCAGGACGCGTATGGGGAGGTGGAGATAAAC 723

Qy        481 GACGGGAAGCTCTACGACGCCTACGTCTCCTACAGCGACTGCCCCGAGGACCGCAAGTTC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        724 GACGGGAAGCTCTACGACGCCTACGTCTCCTACAGCGACTGCCCCGAGGACCGCAAGTTC 783

Qy        541 GTGAACTTCATCCTAAAGCCGCAGCTGGAGCGGCGTCGGGGCTACAAGCTCTTCCTGGAC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        784 GTGAACTTCATCCTAAAGCCGCAGCTGGAGCGGCGTCGGGGCTACAAGCTCTTCCTGGAC 843

Qy        601 GACCGCGACCTCCTGCCGCGCGCTGAGCCCTCCGCCGACCTCTTGGTGAACCTGAGCCGC 660
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        844 GACCGCGACCTCCTGCCCCGCGCTGAGCCCTCCGCCGACCTCTTGGTGAACCTGAGCCGC 903

Qy        661 TGCCGACGCCTCATCGTGGTGCTTTCGGACGCCTTCCTGAGCCGGGCCTGGTGCAGCCAC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        904 TGCCGACGCCTCATCGTGGTGCTTTCGGACGCCTTCCTGAGCCGGGCCTGGTGCAGCCAC 963

Qy        721 AGCTTCCGGGAGGGCCTGTGCCGGCTGCTGGAGCTCACCCGCAGACCCATCTTCATCACC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        964 AGCTTCCGGGAGGGCCTGTGCCGGCTGCTGGAGCTCACCCGCAGACCCATCTTCATCACC 1023

Qy        781 TTCGAGGGCCAGAGGCGCGACCCCGCGCACCCGGCGCTCCGCCTGCTGCGCCAGCACCGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1024 TTCGAGGGCCAGAGGCGCGACCCCGCGCACCCGGCGCTCCGCCTGCTGCGCCAGCACCGC 1083

Qy        841 CACCTGGTGACCTTGCTGCTCTGGAGGCCCGGCTCCGTGACTCCTTCCTCCGATTTTTGG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1084 CACCTGGTGACCTTGCTGCTCTGGAGGCCCGGCTCCGTGACTCCTTCCTCCGATTTTTGG 1143

Qy        901 AAAGAAGTGCAGCTGGCGCTGCCGCGGAAGGTGCGGTACAGGCCGGTGGAAGGAGACCCC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1144 AAAGAAGTGCAGCTGGCGCTGCCGCGGAAGGTGCGGTACAGGCCGGTGGAAGGAGACCCC 1203

Qy        961 CAGACGCAGCTGCAGGACGACAAGGACCCCATGCTGATTCTTCGAGGCCGAGTCCCTGAG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1204 CAGACGCAGCTGCAGGACGACAAGGACCCCATGCTGATTCTTCGAGGCCGAGTCCCTGAG 1263

Qy       1021 GGCCGGGCCCTGGACTCAGAGGTGGACCCGGACCCTGAGGGCGACCTGGGTGTCCGGGGG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1264 GGCCGGGCCCTGGACTCAGAGGTGGACCCGGACCCTGAGGGCGACCTGGGTGTCCGGGGG 1323

Qy       1081 CCTGTTTTTGGAGAGCCATCAGCTCCACCGCACACCAGTGGGGTCTCGCTGGGAGAGAGC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1324 CCTGTTTTTGGAGAGCCATCAGCTCCACCGCACACCAGTGGGGTCTCGCTGGGAGAGAGC 1383

Qy       1141 CGGAGCAGCGAAGTGGACGTCTCGGATCTCGGCTCGCGAAACTACAGTGCCCGCACAGAC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1384 CGGAGCAGCGAAGTGGACGTCTCGGATCTCGGCTCGCGAAACTACAGTGCCCGCACAGAC 1443

Qy       1201 TTCTACTGCCTGGTGTCCAAGGATGATATGTAG 1233
              ||||||||||||||||||||||||||||||||
Db       1444 TTCTACTGCCTGGTGTCCAAGGATGATATGTAG 1476
```

RESULT 9
ADF66744
ID ADF66744 standard; cDNA; 1659 BP.
XX
AC ADF66744;
XX
DT 12-FEB-2004 (first entry)
XX
DE cDNA encoding novel human protein NOV9.
XX
KW cytostatic; hepatotropic; vulnerary; antipsoriatic; osteopathic;
KW antiarthritic; antiatherosclerotic; haemostatic; vasotropic;
KW thrombolytic; antidiabetic; hypotensive; dermatalogical;
KW immunosuppressive; antiinflammatory; immunostimulant; fungicide;
KW virucide; protozoacide; neuroprotective; antirheumatic; antiarthritic;
KW antiasthmatic; antiparkinsonian; nootropic; anticonvulsant;
KW NOVX modulator; cancer; hyperproliferative disease; cirrhosis; keloid;
KW psoriasis; tissue hypertrophy; osteoarthritis;
KW atherosclerotic plaque formation; haemorrhage; ischaemic disease;
KW thrombosis; diabetes mellitus; hypertension; hypothyroidism;
KW immune deficiency; severe combined immunodeficiency; SCID; infection;
KW malaria; candidiasis; autoimmune disorder; connective tissue disease;
KW multiple sclerosis; systemic lupus, erythematosus; rheumatoid arthritis;
KW autoimmune pulmonary inflammation; Guillain-Barre syndrome;
KW autoimmune thyroiditis; insulin dependent diabetes mellitus;
KW myasthenia gravis; graft- versus-host disease;
KW autoimmune inflammatory eye disease; asthma; haematopoiesis;
KW tissue regeneration; wound healing; tissue repair; burn; incision; ulcer;
KW periodontal disease; Alzhemier's disease; Parkinson's disease;
KW Huntington's disease; amyotrophic lateral sclerosis; Shy-Drager syndrome;
KW human; gene; ss.
XX
OS Homo sapiens.
XX
PN US2003199103-A1.
XX
PD 23-OCT-2003.
XX
PF 15-OCT-2001; 2001US-00977639.
XX
PR 03-JUN-1999; 99US-0137322P.
PR 16-MAR-2000; 2000US-0189810P.
PR 22-MAR-2000; 2000US-0191158P.
PR 30-MAR-2000; 2000US-0193086P.
PR 03-MAY-2000; 2000US-0201388P.
PR 31-MAY-2000; 2000US-00584411.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Fernandes E, Herrman J, Vernet C;
XX
DR WPI; 2004-021196/02.
DR P-PSDB; ADF66745.
XX
PT Novel substantially NOVX polypeptide useful for diagnosing, preventing
PT and treating diseases e.g., cancer, multiple sclerosis, systemic lupus
PT erythematous.
XX
PS Disclosure; SEQ ID NO 17; 165pp; English.
XX
CC The invention describes a substantially purified polypeptide (I) having
CC amino acid sequence chosen from a fully defined NOVX sequence (S1) of 708
CC amino acids as given in the specification, or polypeptide having one or
CC more conservative amino acid substitutions of (S1), or mutant or variant
CC of (S1). (I) having (S1) is useful for diagnosing a pathological
CC condition associated with (I) or its activity in a subject e.g. cancer.
CC (I) useful in treatment of cancer, hyperproliferative diseases,
CC cirrhosis, keloid, psoriasis, tissue hypertrophy, osteoarthritis,
CC atherosclerotic plaque formation, haemorrhage, ischaemic heart or renal
CC disease, thrombosis, diabetes mellitus, hypertension, hypothyroidism. (I)
CC is useful in treatment of various immune deficiencies and disorders such
CC as severe combined immunodeficiency (SCID), bacterial infection, viral
CC infection such as herpes viral infection, protozoan infection such as
CC malaria, fungal infection such as candidiasis. (I) is also useful in
CC treating autoimmune disorders such as connective tissue disease, multiple
CC sclerosis, systemic lupus, erythematosus, rheumatoid arthritis,
CC autoimmune pulmonary inflammation, Guillain-Barre syndrome, autoimmune
CC thyroiditis, insulin dependent diabetes mellitus, myasthenia gravis,
CC graft- versus-host disease and autoimmune inflammatory eye disease and
CC asthma. (I) useful in regulation of haematopoiesis, regeneration and
CC tissue growth of bone, cartilage, tendon, ligament and useful for wound
CC healing and tissue repair. (I) is also useful in treatment of burns,
CC incisions and ulcers. (I) also useful in treatment of periodontal
CC disease, Alzhemier's disease, Parkinson's disease, Huntington's disease,
CC amyotrophic lateral sclerosis, and Shy-Drager syndrome. (I) has effective

CC antitumour and antiinflammatory activity. This sequence encodes a novel
CC human NOVX protein.
XX
SQ Sequence 1659 BP; 296 A; 551 C; 504 G; 308 T; 0 U; 0 Other;

Query Match 99.9%; Score 1231.4; DB 12; Length 1659;
Best Local Similarity 99.9%; Pred. No. 2.3e-236;
Matches 1232; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 ATGCCAGGTGTCTGTGATAGGGCCCCTGACTTCCTCTCCCCGTCTGAAGACCAGGTGCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 ATGCCAGGTGTCTGTGATAGGGCCCCTGACTTCCTCTCCCCGTCTGAAGACCAGGTGCTG 303

Qy         61 AGGCCTGCCTTGGGCAGCTCAGTGGCTCTGAACTGCACGGCTTGGGTAGTCTCTGGGCCC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        304 AGGCCTGCCTTGGGCAGCTCAGTGGCTCTGAACTGCACGGCTTGGGTAGTCTCTGGGCCC 363

Qy        121 CACTGCTCCCTGCCTTCAGTCCAGTGGCTGAAAGACGGGCTTCCATTGGGAATTGGGGGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        364 CACTGCTCCCTGCCTTCAGTCCAGTGGCTGAAAGACGGGCTTCCATTGGGAATTGGGGGC 423

Qy        181 CACTACAGCCTCCACGAGTACTCCTGGGTCAAGGCCAACCTGTCAGAGGTGCTTGTGTCC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        424 CACTACAGCCTCCACGAGTACTCCTGGGTCAAGGCCAACCTGTCAGAGGTGCTTGTGTCC 483

Qy        241 AGTGTCCTGGGGGTCAACGTGACCAGCACTGAAGTCTATGGGGCCTTCACCTGCTCCATC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        484 AGTGTCCTGGGGGTCAACGTGACCAGCACTGAAGTCTATGGGGCCTTCACCTGCTCCATC 543

Qy        301 CAGAACATCAGCTTCTCCTCCTTCACTCTTCAGAGAGCTGGCCCTACAAGCCACGTGGCT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        544 CAGAACATCAGCTTCTCCTCCTTCACTCTTCAGAGAGCTGGCCCTACAAGCCACGTGGCT 603

Qy        361 GCGGTGCTGGCCTCCCTCCTGGTCCTGCTGGCCCTGCTGCTGGCCGCCCTGCTCTATGTC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        604 GCGGTGCTGGCCTCCCTCCTGGTCCTGCTGGCCCTGCTGCTGGCCGCCCTGCTCTATGTC 663

Qy        421 AAGTGCCGTCTCAACGTGCTGCTCTGGTACCAGGACGCGTATGGGGAGGTGGAGATAAAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        664 AAGTGCCGTCTCAACGTGCTGCTCTGGTACCAGGACGCGTATGGGGAGGTGGAGATAAAC 723

Qy        481 GACGGGAAGCTCTACGACGCCTACGTCTCCTACAGCGACTGCCCCGAGGACCGCAAGTTC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        724 GACGGGAAGCTCTACGACGCCTACGTCTCCTACAGCGACTGCCCCGAGGACCGCAAGTTC 783

Qy        541 GTGAACTTCATCCTAAAGCCGCAGCTGGAGCGGCGTCGGGGCTACAAGCTCTTCCTGGAC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        784 GTGAACTTCATCCTAAAGCCGCAGCTGGAGCGGCGTCGGGGCTACAAGCTCTTCCTGGAC 843

Qy        601 GACCGCGACCTCCTGCCGCGCGCTGAGCCCTCCGCCGACCTCTTGGTGAACCTGAGCCGC 660
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        844 GACCGCGACCTCCTGCCCCGCGCTGAGCCCTCCGCCGACCTCTTGGTGAACCTGAGCCGC 903

Qy        661 TGCCGACGCCTCATCGTGGTGCTTTCGGACGCCTTCCTGAGCCGGGCCTGGTGCAGCCAC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        904 TGCCGACGCCTCATCGTGGTGCTTTCGGACGCCTTCCTGAGCCGGGCCTGGTGCAGCCAC 963

Qy        721 AGCTTCCGGGAGGGCCTGTGCCGGCTGCTGGAGCTCACCCGCAGACCCATCTTCATCACC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        964 AGCTTCCGGGAGGGCCTGTGCCGGCTGCTGGAGCTCACCCGCAGACCCATCTTCATCACC 1023

Qy        781 TTCGAGGGCCAGAGGCGCGACCCCGCGCACCCGGCGCTCCGCCTGCTGCGCCAGCACCGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1024 TTCGAGGGCCAGAGGCGCGACCCCGCGCACCCGGCGCTCCGCCTGCTGCGCCAGCACCGC 1083

Qy        841 CACCTGGTGACCTTGCTGCTCTGGAGGCCCGGCTCCGTGACTCCTTCCTCCGATTTTTGG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1084 CACCTGGTGACCTTGCTGCTCTGGAGGCCCGGCTCCGTGACTCCTTCCTCCGATTTTTGG 1143

Qy        901 AAAGAAGTGCAGCTGGCGCTGCCGCGGAAGGTGCGGTACAGGCCGGTGGAAGGAGACCCC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1144 AAAGAAGTGCAGCTGGCGCTGCCGCGGAAGGTGCGGTACAGGCCGGTGGAAGGAGACCCC 1203

Qy        961 CAGACGCAGCTGCAGGACGACAAGGACCCCATGCTGATTCTTCGAGGCCGAGTCCCTGAG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1204 CAGACGCAGCTGCAGGACGACAAGGACCCCATGCTGATTCTTCGAGGCCGAGTCCCTGAG 1263

Qy       1021 GGCCGGGCCCTGGACTCAGAGGTGGACCCGGACCCTGAGGGCGACCTGGGTGTCCGGGGG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1264 GGCCGGGCCCTGGACTCAGAGGTGGACCCGGACCCTGAGGGCGACCTGGGTGTCCGGGGG 1323

Qy       1081 CCTGTTTTTGGAGAGCCATCAGCTCCACCGCACACCAGTGGGGTCTCGCTGGGAGAGAGC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1324 CCTGTTTTTGGAGAGCCATCAGCTCCACCGCACACCAGTGGGGTCTCGCTGGGAGAGAGC 1383

Qy       1141 CGGAGCAGCGAAGTGGACGTCTCGGATCTCGGCTCGCGAAACTACAGTGCCCGCACAGAC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1384 CGGAGCAGCGAAGTGGACGTCTCGGATCTCGGCTCGCGAAACTACAGTGCCCGCACAGAC 1443

Qy       1201 TTCTACTGCCTGGTGTCCAAGGATGATATGTAG 1233
              ||||||||||||||||||||||||||||||||
Db       1444 TTCTACTGCCTGGTGTCCAAGGATGATATGTAG 1476
```

RESULT 10
ADO60254
ID ADO60254 standard; DNA; 1659 BP.
XX
AC ADO60254;
XX
DT 15-JUL-2004 (first entry)
XX
DE Human NOV9 DNA.
XX
KW Human; NOVX protein; cancer; hyperproliferative disease; cirrhosis;
KW keloid; psoriasis; tissue hypertrophy; osteoarthritis;
KW atherosclerotic plaque formation; haemorrhage; ischaemic heart disease;
KW renal disease; thrombosis; diabetes mellitus; hypertension;
KW hypothyroidism; severe combined immunodeficiency; SCID; infection;
KW malaria; candidiasis; autoimmune disorder; connective tissue disease;
KW multiple sclerosis; systemic lupus erythematosus; rheumatoid arthritis;
KW autoimmune pulmonary inflammation; Guillain-Barre syndrome;
KW autoimmune thyroiditis; myasthenia gravis; graft- versus-host disease;
KW autoimmune inflammatory eye disease; asthma; burn; incision; ulcer;
KW periodontal disease; Alzhemier's disease; Parkinson's disease;
KW Huntington's disease; amyotrophic lateral sclerosis; Shy-Drager syndrome;
KW haematopoiesis; wound healing; tissue repair; antitumor;
KW antiinflammatory; gene; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 244. .1476
FT /*tag= a
FT /product= "NOVX protein"
XX
PN US2003134430-A1.
XX
PD 17-JUL-2003.
XX
PF 15-OCT-2001; 2001US-00977751.
XX
PR 03-JUN-1999; 99US-0137322P.
PR 16-MAR-2000; 2000US-0189810P.
PR 22-MAR-2000; 2000US-0191158P.
PR 30-MAR-2000; 2000US-0193086P.
PR 03-MAY-2000; 2000US-0201388P.
PR 31-MAY-2000; 2000US-00584411.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Fernandes E, Herrman J, Vernet C;
XX
DR WPI; 2004-068928/07.
DR P-PSDB; ADO60255.
XX

```
PT   Novel substantially purified NOVX polypeptide for treating severe
PT   combined immunodeficiency, candidiasis, cancer, asthma, multiple
PT   sclerosis, systemic lupus erythematosus.
XX
PS   Disclosure; SEQ ID NO 17; 155pp; English.
XX
CC   The invention relates to human NOVX polypeptides and polynucleotides.
CC   NOVX sequences are useful in the treatment of cancer, hyperproliferative
CC   diseases, cirrhosis, keloid, psoriasis, tissue hypertrophy,
CC   osteoarthritis, atherosclerotic plaque formation, haemorrhage, ischaemic
CC   heart or renal disease, thrombosis, diabetes mellitus, hypertension,
CC   hypothyroidism, asthma, burns, incisions, ulcers, periodontal disease,
CC   Alzhemier's disease, Parkinson's disease, Huntington's disease,
CC   amyotrophic lateral sclerosis, Shy-Drager syndrome, immune deficiencies
CC   and disorders such as severe combined immunodeficiency (SCID), bacterial
CC   infection, viral infection e.g. herpes viral infection, protozoan
CC   infection e.g. malaria, fungal infection e.g. candidiasis, autoimmune
CC   disorders such as connective tissue disease, multiple sclerosis, systemic
CC   lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary
CC   inflammation, Guillain-Barre syndrome, autoimmune thyroiditis, insulin
CC   dependent diabetes mellitus, myasthenia gravis, graft-versus-host disease
CC   and autoimmune inflammatory eye disease. The invention is useful in
CC   regeneration and tissue growth of bone, cartilage, tendon, ligament,
CC   haematopoiesis regulation, wound healing and tissue repair. Sequences of
CC   the invention also exhibit antitumor and antiinflammatory activities. The
CC   present sequence is human NOVX DNA.
XX
SQ   Sequence 1659 BP; 296 A; 551 C; 504 G; 308 T; 0 U; 0 Other;

  Query Match             99.9%;  Score 1231.4;  DB 12;  Length 1659;
  Best Local Similarity   99.9%;  Pred. No. 2.3e-236;
  Matches 1232;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 ATGCCAGGTGTCTGTGATAGGGCCCCTGACTTCCTCTCCCCGTCTGAAGACCAGGTGCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 ATGCCAGGTGTCTGTGATAGGGCCCCTGACTTCCTCTCCCCGTCTGAAGACCAGGTGCTG 303

Qy         61 AGGCCTGCCTTGGGCAGCTCAGTGGCTCTGAACTGCACGGCTTGGGTAGTCTCTGGGCCC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        304 AGGCCTGCCTTGGGCAGCTCAGTGGCTCTGAACTGCACGGCTTGGGTAGTCTCTGGGCCC 363

Qy        121 CACTGCTCCCTGCCTTCAGTCCAGTGGCTGAAAGACGGGCTTCCATTGGGAATTGGGGGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        364 CACTGCTCCCTGCCTTCAGTCCAGTGGCTGAAAGACGGGCTTCCATTGGGAATTGGGGGC 423

Qy        181 CACTACAGCCTCCACGAGTACTCCTGGGTCAAGGCCAACCTGTCAGAGGTGCTTGTGTCC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        424 CACTACAGCCTCCACGAGTACTCCTGGGTCAAGGCCAACCTGTCAGAGGTGCTTGTGTCC 483

Qy        241 AGTGTCCTGGGGGTCAACGTGACCAGCACTGAAGTCTATGGGGCCTTCACCTGCTCCATC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        484 AGTGTCCTGGGGGTCAACGTGACCAGCACTGAAGTCTATGGGGCCTTCACCTGCTCCATC 543

Qy        301 CAGAACATCAGCTTCTCCTCCTTCACTCTTCAGAGAGCTGGCCCTACAAGCCACGTGGCT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        544 CAGAACATCAGCTTCTCCTCCTTCACTCTTCAGAGAGCTGGCCCTACAAGCCACGTGGCT 603

Qy        361 GCGGTGCTGGCCTCCCTCCTGGTCCTGCTGGCCCTGCTGCTGGCCGCCCTGCTCTATGTC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        604 GCGGTGCTGGCCTCCCTCCTGGTCCTGCTGGCCCTGCTGCTGGCCGCCCTGCTCTATGTC 663

Qy        421 AAGTGCCGTCTCAACGTGCTGCTCTGGTACCAGGACGCGTATGGGGAGGTGGAGATAAAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        664 AAGTGCCGTCTCAACGTGCTGCTCTGGTACCAGGACGCGTATGGGGAGGTGGAGATAAAC 723

Qy        481 GACGGGAAGCTCTACGACGCCTACGTCTCCTACAGCGACTGCCCCGAGGACCGCAAGTTC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        724 GACGGGAAGCTCTACGACGCCTACGTCTCCTACAGCGACTGCCCCGAGGACCGCAAGTTC 783

Qy        541 GTGAACTTCATCCTAAAGCCGCAGCTGGAGCGGCGTCGGGGCTACAAGCTCTTCCTGGAC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        784 GTGAACTTCATCCTAAAGCCGCAGCTGGAGCGGCGTCGGGGCTACAAGCTCTTCCTGGAC 843

Qy        601 GACCGCGACCTCCTGCCGCGCGCTGAGCCCTCCGCCGACCTCTTGGTGAACCTGAGCCGC 660
              ||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db        844 GACCGCGACCTCCTGCCCCGCGCTGAGCCCTCCGCCGACCTCTTGGTGAACCTGAGCCGC 903

Qy        661 TGCCGACGCCTCATCGTGGTGCTTTCGGACGCCTTCCTGAGCCGGGCCTGGTGCAGCCAC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        904 TGCCGACGCCTCATCGTGGTGCTTTCGGACGCCTTCCTGAGCCGGGCCTGGTGCAGCCAC 963

Qy        721 AGCTTCCGGGAGGGCCTGTGCCGGCTGCTGGAGCTCACCCGCAGACCCATCTTCATCACC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        964 AGCTTCCGGGAGGGCCTGTGCCGGCTGCTGGAGCTCACCCGCAGACCCATCTTCATCACC 1023

Qy        781 TTCGAGGGCCAGAGGCGCGACCCCGCGCACCCGGCGCTCCGCCTGCTGCGCCAGCACCGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1024 TTCGAGGGCCAGAGGCGCGACCCCGCGCACCCGGCGCTCCGCCTGCTGCGCCAGCACCGC 1083

Qy        841 CACCTGGTGACCTTGCTGCTCTGGAGGCCCGGCTCCGTGACTCCTTCCTCCGATTTTTGG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1084 CACCTGGTGACCTTGCTGCTCTGGAGGCCCGGCTCCGTGACTCCTTCCTCCGATTTTTGG 1143

Qy        901 AAAGAAGTGCAGCTGGCGCTGCCGCGGAAGGTGCGGTACAGGCCGGTGGAAGGAGACCCC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1144 AAAGAAGTGCAGCTGGCGCTGCCGCGGAAGGTGCGGTACAGGCCGGTGGAAGGAGACCCC 1203

Qy        961 CAGACGCAGCTGCAGGACGACAAGGACCCCATGCTGATTCTTCGAGGCCGAGTCCCTGAG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1204 CAGACGCAGCTGCAGGACGACAAGGACCCCATGCTGATTCTTCGAGGCCGAGTCCCTGAG 1263

Qy       1021 GGCCGGGCCCTGGACTCAGAGGTGGACCCGGACCCTGAGGGCGACCTGGGTGTCCGGGGG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1264 GGCCGGGCCCTGGACTCAGAGGTGGACCCGGACCCTGAGGGCGACCTGGGTGTCCGGGGG 1323

Qy       1081 CCTGTTTTTGGAGAGCCATCAGCTCCACCGCACACCAGTGGGGTCTCGCTGGGAGAGAGC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1324 CCTGTTTTTGGAGAGCCATCAGCTCCACCGCACACCAGTGGGGTCTCGCTGGGAGAGAGC 1383

Qy       1141 CGGAGCAGCGAAGTGGACGTCTCGGATCTCGGCTCGCGAAACTACAGTGCCCGCACAGAC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1384 CGGAGCAGCGAAGTGGACGTCTCGGATCTCGGCTCGCGAAACTACAGTGCCCGCACAGAC 1443

Qy       1201 TTCTACTGCCTGGTGTCCAAGGATGATATGTAG 1233
              |||||||||||||||||||||||||||||||||
Db       1444 TTCTACTGCCTGGTGTCCAAGGATGATATGTAG 1476


RESULT 11
ADI19781
ID   ADI19781 standard; DNA; 1659 BP.
XX
AC   ADI19781;
XX
DT   22-APR-2004  (first entry)
XX
DE   Human NOV9 DNA.
XX
KW   Secreted protein; NOVX; diagnosis; metabolic disorder; diabetes; obesity;
KW   infection; anorexia; cancer; cardiovascular disease; hypertension;
KW   atherosclerosis; neurodegenerative disorder; Alzheimer's disease;
KW   Parkinson's disease; epilepsy; immune disorder; osteoarthritis;
KW   haematopoietic disorder; inflammatory skin disorder; asthma;
KW   dyslipidemia; neurogenesis; cell differentiation; cell proliferation;
KW   haematopoiesis; wound healing; angiogenesis; chromosome mapping;
KW   tissue typing; preventive medicine; pharmacogenomic; gene therapy;
KW   anorectic; cardiant; virucide; antibacterial; fungicide; protozoacide;
KW   nootropic; neuroprotective; dermatological; human; gene; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             244. .1476
FT                   /*tag=  a
FT                   /product= "Human NOV protein"
FT                   /transl_except= (pos352. .354, aa:Val)
```

XX
PN US2004002134-A1.
XX
PD 01-JAN-2004.
XX
PF 15-OCT-2001; 2001US-00977819.
XX
PR 03-JUN-1999; 99US-0137322P.
PR 16-MAR-2000; 2000US-0189810P.
PR 22-MAR-2000; 2000US-0191158P.
PR 30-MAR-2000; 2000US-0193086P.
PR 03-MAY-2000; 2000US-0201388P.
PR 31-MAY-2000; 2000US-00584411.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Fernandes ER, Herrman JL, Vernet CAM;
XX
DR WPI; 2004-070737/07.
DR P-PSDB; ADI19782.
XX
PT New NOVX nucleic acids encoding human KIAA0768 protein-like and human
PT protein PRO-228 polypeptides, useful for treating NOVX-associated
PT disorders.
XX
PS Disclosure; SEQ ID NO 17; 95pp; English.
XX
CC The present invention is based in part on the discovery of novel secreted
CC and membrane-bound polypeptides and their encoding polynucleotides. The
CC nucleic acids and polypeptides are collectively referred as NOVX. The
CC invention is useful for treating, preventing and diagnosing diseases such
CC as metabolic disorders, diabetes, obesity, infectious diseases such as
CC viral, bacterial, fungal, helminthic and protozoal infections, anorexia,
CC cancer, cardiovascular diseases such as hypertension and atherosclerosis,
CC neurodegenerative disorders, Alzheimer's disease, Parkinson's disease,
CC epilepsy, immune disorders such as osteoarthritis, haematopoietic
CC disorders, inflammatory skin disorders, asthma and various dyslipidemias.
CC The invention is also useful as targets for the identification of small
CC molecules that modulate or inhibit e.g. neurogenesis, cell
CC differentiation, cell proliferation, haematopoiesis, wound healing and
CC angiogenesis, as hybridisation probes, in chromosome mapping, tissue
CC typing, preventive medicine and pharmacogenomics. The invention is also
CC useful in gene therapy. The present sequence is human NOV DNA.
XX
SQ Sequence 1659 BP; 295 A; 552 C; 502 G; 308 T; 0 U; 2 Other;

Query Match 99.8%; Score 1230.4; DB 12; Length 1659;
Best Local Similarity 99.8%; Pred. No. 3.7e-236;
Matches 1231; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 ATGCCAGGTGTCTGTGATAGGGCCCCTGACTTCCTCTCCCCGTCTGAAGACCAGGTGCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        244 ATGCCAGGTGTCTGTGATAGGGCCCCTGACTTCCTCTCCCCGTCTGAAGACCAGGTGCTG 303

Qy         61 AGGCCTGCCTTGGGCAGCTCAGTGGCTCTGAACTGCACGGCTTGGGTAGTCTCTGGGCCC 120
              |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db        304 AGGCCTGCCTTGGGCAGCTCAGTGGCTCTGAACTGCACGGCTTGGGTANTCTCTGGGCCC 363

Qy        121 CACTGCTCCCTGCCTTCAGTCCAGTGGCTGAAAGACGGGCTTCCATTGGGAATTGGGGGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        364 CACTGCTCCCTGCCTTCAGTCCAGTGGCTGAAAGACGGGCTTCCATTGGGAATTGGGGGC 423

Qy        181 CACTACAGCCTCCACGAGTACTCCTGGGTCAAGGCCAACCTGTCAGAGGTGCTTGTGTCC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        424 CACTACAGCCTCCACGAGTACTCCTGGGTCAAGGCCAACCTGTCAGAGGTGCTTGTGTCC 483

Qy        241 AGTGTCCTGGGGGTCAACGTGACCAGCACTGAAGTCTATGGGGCCTTCACCTGCTCCATC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        484 AGTGTCCTGGGGGTCAACGTGACCAGCACTGAAGTCTATGGGGCCTTCACCTGCTCCATC 543

Qy        301 CAGAACATCAGCTTCTCCTCCTTCACTCTTCAGAGAGCTGGCCCTACAAGCCACGTGGCT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        544 CAGAACATCAGCTTCTCCTCCTTCACTCTTCAGAGAGCTGGCCCTACAAGCCACGTGGCT 603

Qy        361 GCGGTGCTGGCCTCCCTCCTGGTCCTGCTGGCCCTGCTGCTGGCCGCCCTGCTCTATGTC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        604 GCGGTGCTGGCCTCCCTCCTGGTCCTGCTGGCCCTGCTGCTGGCCGCCCTGCTCTATGTC 663

Qy        421 AAGTGCCGTCTCAACGTGCTGCTCTGGTACCAGGACGCGTATGGGGAGGTGGAGATAAAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        664 AAGTGCCGTCTCAACGTGCTGCTCTGGTACCAGGACGCGTATGGGGAGGTGGAGATAAAC 723

Qy        481 GACGGGAAGCTCTACGACGCCTACGTCTCCTACAGCGACTGCCCCGAGGACCGCAAGTTC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        724 GACGGGAAGCTCTACGACGCCTACGTCTCCTACAGCGACTGCCCCGAGGACCGCAAGTTC 783

Qy        541 GTGAACTTCATCCTAAAGCCGCAGCTGGAGCGGCGTCGGGGCTACAAGCTCTTCCTGGAC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        784 GTGAACTTCATCCTAAAGCCGCAGCTGGAGCGGCGTCGGGGCTACAAGCTCTTCCTGGAC 843

Qy        601 GACCGCGACCTCCTGCCGCGCGCTGAGCCCTCCGCCGACCTCTTGGTGAACCTGAGCCGC 660
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        844 GACCGCGACCTCCTGCCCCGCGCTGAGCCCTCCGCCGACCTCTTGGTGAACCTGAGCCGC 903

Qy        661 TGCCGACGCCTCATCGTGGTGCTTTCGGACGCCTTCCTGAGCCGGGCCTGGTGCAGCCAC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        904 TGCCGACGCCTCATCGTGGTGCTTTCGGACGCCTTCCTGAGCCGGGCCTGGTGCAGCCAC 963

Qy        721 AGCTTCCGGGAGGGCCTGTGCCGGCTGCTGGAGCTCACCCGCAGACCCATCTTCATCACC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        964 AGCTTCCGGGAGGGCCTGTGCCGGCTGCTGGAGCTCACCCGCAGACCCATCTTCATCACC 1023

Qy        781 TTCGAGGGCCAGAGGCGCGACCCCGCGCACCCGGCGCTCCGCCTGCTGCGCCAGCACCGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1024 TTCGAGGGCCAGAGGCGCGACCCCGCGCACCCGGCGCTCCGCCTGCTGCGCCAGCACCGC 1083

Qy        841 CACCTGGTGACCTTGCTGCTCTGGAGGCCCGGCTCCGTGACTCCTTCCTCCGATTTTTGG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1084 CACCTGGTGACCTTGCTGCTCTGGAGGCCCGGCTCCGTGACTCCTTCCTCCGATTTTTGG 1143

Qy        901 AAAGAAGTGCAGCTGGCGCTGCCGCGGAAGGTGCGGTACAGGCCGGTGGAAGGAGACCCC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1144 AAAGAAGTGCAGCTGGCGCTGCCGCGGAAGGTGCGGTACAGGCCGGTGGAAGGAGACCCC 1203

Qy        961 CAGACGCAGCTGCAGGACGACAAGGACCCCATGCTGATTCTTCGAGGCCGAGTCCCTGAG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1204 CAGACGCAGCTGCAGGACGACAAGGACCCCATGCTGATTCTTCGAGGCCGAGTCCCTGAG 1263

Qy       1021 GGCCGGGCCCTGGACTCAGAGGTGGACCCGGACCCTGAGGGCGACCTGGGTGTCCGGGGG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1264 GGCCGGGCCCTGGACTCAGAGGTGGACCCGGACCCTGAGGGCGACCTGGGTGTCCGGGGG 1323

Qy       1081 CCTGTTTTTGGAGAGCCATCAGCTCCACCGCACACCAGTGGGGTCTCGCTGGGAGAGAGC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1324 CCTGTTTTTGGAGAGCCATCAGCTCCACCGCACACCAGTGGGGTCTCGCTGGGAGAGAGC 1383

Qy       1141 CGGAGCAGCGAAGTGGACGTCTCGGATCTCGGCTCGCGAAACTACAGTGCCCGCACAGAC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1384 CGGAGCAGCGAAGTGGACGTCTCGGATCTCGGCTCGCGAAACTACAGTGCCCGCACAGAC 1443

Qy       1201 TTCTACTGCCTGGTGTCCAAGGATGATATGTAG 1233
              |||||||||||||||||||||||||||||||||
Db       1444 TTCTACTGCCTGGTGTCCAAGGATGATATGTAG 1476
```

RESULT 12
ADP46448
ID ADP46448 standard; DNA; 1833 BP.
XX
AC ADP46448;
XX
DT 09-SEP-2004 (first entry)
XX
DE Human colon specific nucleic acid SEQ ID NO:24.
XX
KW ds; gene; human; colon specific protein; CSP; CSNA; cytostatic; vaccine;

```
KW   gene therapy; colon specific nucleic acid; cancer; colon;
KW   chromosome 11p15.5.
XX
OS   Homo sapiens.
XX
PN   WO2004050860-A2.
XX
PD   17-JUN-2004.
XX
PF   04-DEC-2003; 2003WO-US040063.
XX
PR   04-DEC-2002; 2002US-0431132P.
PR   04-DEC-2002; 2002US-0431144P.
XX
PA   (DIAD-) DIADEXUS INC.
XX
PI   Macina RA,  Turner LR,  Sun Y;
XX
DR   WPI; 2004-480623/45.
XX
PT   Novel colon specific protein derived from normal and neoplastic colon
PT   cell, useful as vaccine in treating colon cancer and in identifying,
PT   diagnosing, monitoring, staging, imaging colon cancer and non-cancerous
PT   disease state in colon.
XX
PS   Claim 1; SEQ ID NO 24; 547pp; English.
XX
CC   The invention relates to a novel colon specific protein (CSP) (I)and the
CC   nucleic acid (CSNA) encoding it. A CSP of the invention has cytostatic
CC   activity. The protein is useful as a vaccine, and the nucleic acid may
CC   have a use in gene therapy. A CSP is useful for determining the presence
CC   of a colon specific protein in a sample. A CSNA is useful for determining
CC   the presence of a colon specific nucleic acid (CSNA) in a sample. The CSP
CC   and CSNA are useful for diagnosing or monitoring the presence and
CC   metastases of colon cancer in a patient. A method of the invention is
CC   useful for treating a patient with colon cancer. A CSP and CSNA are
CC   useful in identifying, diagnosing, monitoring, staging, imaging colon
CC   cancer and non-cancerous disease state in colon, and as a vaccine for
CC   treating colon cancer and non-cancerous disease states in colon. The
CC   present sequence represents a CSNA of the invention.
XX
SQ   Sequence 1833 BP; 330 A; 610 C; 564 G; 329 T; 0 U; 0 Other;

  Query Match             93.8%;  Score 1156;  DB 12;  Length 1833;
  Best Local Similarity   94.8%;  Pred. No. 2.6e-221;
  Matches 1233;  Conservative    0;  Mismatches    0;  Indels   67;  Gaps    1;

Qy          1 ATGCCAGGTGTCTGTGATAGGGCCCCTGACTTCCTCTCCCCGTCTGAAGACCAGGTGCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        353 ATGCCAGGTGTCTGTGATAGGGCCCCTGACTTCCTCTCCCCGTCTGAAGACCAGGTGCTG 412

Qy         61 AGGCCTGCCTTGGGCAGCTCAGTGGCTCTGAACTGCACGGCTTGGGTAGTCTCTGGGCCC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        413 AGGCCTGCCTTGGGCAGCTCAGTGGCTCTGAACTGCACGGCTTGGGTAGTCTCTGGGCCC 472

Qy        121 CACTGCTCCCTGCCTTCAGTCCAGTGGCTGAAAGACGGGCTTCCATTGGGAATTGGGGGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        473 CACTGCTCCCTGCCTTCAGTCCAGTGGCTGAAAGACGGGCTTCCATTGGGAATTGGGGGC 532

Qy        181 CACTACAGCCTCCACGAGTACTCCTGGGTCAAGGCCAACCTGTCAGAGGTGCTTGTGTCC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        533 CACTACAGCCTCCACGAGTACTCCTGGGTCAAGGCCAACCTGTCAGAGGTGCTTGTGTCC 592

Qy        241 AGTGTCCTGGGGGTCAACGTGACCAGCACTGAAGTCTATGGGGCCTTCACCTGCTCCATC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        593 AGTGTCCTGGGGGTCAACGTGACCAGCACTGAAGTCTATGGGGCCTTCACCTGCTCCATC 652

Qy        301 CAGAACATCAGCTTCTCCTCCTTCACTCTTCAGAGAGCTGGCCCTACAAGCCACGTGGCT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        653 CAGAACATCAGCTTCTCCTCCTTCACTCTTCAGAGAGCTGGCCCTACAAGCCACGTGGCT 712

Qy        361 GCGGTGCTGGCCTCCCTCCTGGTCCTGCTGGCCCTGCTGCTGGCCGCCCTGCTCTATGTC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        713 GCGGTGCTGGCCTCCCTCCTGGTCCTGCTGGCCCTGCTGCTGGCCGCCCTGCTCTATGTC 772

Qy        421 AAGTGCCGTCTCAACGTGCTGCTCTGGTACCAGGACGCGTATGGGGAGGTGGAGATAAAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        773 AAGTGCCGTCTCAACGTGCTGCTCTGGTACCAGGACGCGTATGGGGAGGTGGAGATAAAC 832

Qy        481 GACGGGAAGCTCTACGACGCCTACGTCTCCTACAGCGACTGCCCCGAGGACCGCAAGTTC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        833 GACGGGAAGCTCTACGACGCCTACGTCTCCTACAGCGACTGCCCCGAGGACCGCAAGTTC 892

Qy        541 GTGAACTTCATCCTAAAGCCGCAGCTGGAGCGGCGTCGGGGCTACAAGCTCTTCCTGGAC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        893 GTGAACTTCATCCTAAAGCCGCAGCTGGAGCGGCGTCGGGGCTACAAGCTCTTCCTGGAC 952

Qy        601 GACCGCGACCTCCTGCCGCGCGCTGAGCCCTCCGCCGACCTCTTGGTGAACCTGAGCCGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        953 GACCGCGACCTCCTGCCGCGCGCTGAGCCCTCCGCCGACCTCTTGGTGAACCTGAGCCGC 1012

Qy        661 TGCCGACGCCTCATCGTGGTGCTTTCGGACGCCTTCCTGAGCCGGGCCTGGTGCAGCCAC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1013 TGCCGACGCCTCATCGTGGTGCTTTCGGACGCCTTCCTGAGCCGGGCCTGGTGCAGCCAC 1072

Qy        721 AGCTTCC------------------------------------------------------ 727
              |||||||
Db       1073 AGCTTCCGGTGGGTCCCGCGCGGGGTTGGGTGGGCCCCAGCGTACACCCACCCCCCTGAC 1132

Qy        728 --------------GGGAGGGCCTGTGCCGGCTGCTGGAGCTCACCCGCAGACCCATCTT 773
                          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1133 GGTCCCGCCCCGCAGGGAGGGCCTGTGCCGGCTGCTGGAGCTCACCCGCAGACCCATCTT 1192

Qy        774 CATCACCTTCGAGGGCCAGAGGCGCGACCCCGCGCACCCGGCGCTCCGCCTGCTGCGCCA 833
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1193 CATCACCTTCGAGGGCCAGAGGCGCGACCCCGCGCACCCGGCGCTCCGCCTGCTGCGCCA 1252

Qy        834 GCACCGCCACCTGGTGACCTTGCTGCTCTGGAGGCCCGGCTCCGTGACTCCTTCCTCCGA 893
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1253 GCACCGCCACCTGGTGACCTTGCTGCTCTGGAGGCCCGGCTCCGTGACTCCTTCCTCCGA 1312

Qy        894 TTTTTGGAAAGAAGTGCAGCTGGCGCTGCCGCGGAAGGTGCGGTACAGGCCGGTGGAAGG 953
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1313 TTTTTGGAAAGAAGTGCAGCTGGCGCTGCCGCGGAAGGTGCGGTACAGGCCGGTGGAAGG 1372

Qy        954 AGACCCCCAGACGCAGCTGCAGGACGACAAGGACCCCATGCTGATTCTTCGAGGCCGAGT 1013
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1373 AGACCCCCAGACGCAGCTGCAGGACGACAAGGACCCCATGCTGATTCTTCGAGGCCGAGT 1432

Qy       1014 CCCTGAGGGCCGGGCCCTGGACTCAGAGGTGGACCCGGACCCTGAGGGCGACCTGGGTGT 1073
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1433 CCCTGAGGGCCGGGCCCTGGACTCAGAGGTGGACCCGGACCCTGAGGGCGACCTGGGTGT 1492

Qy       1074 CCGGGGGCCTGTTTTTGGAGAGCCATCAGCTCCACCGCACACCAGTGGGGTCTCGCTGGG 1133
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1493 CCGGGGGCCTGTTTTTGGAGAGCCATCAGCTCCACCGCACACCAGTGGGGTCTCGCTGGG 1552

Qy       1134 AGAGAGCCGGAGCAGCGAAGTGGACGTCTCGGATCTCGGCTCGCGAAACTACAGTGCCCG 1193
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1553 AGAGAGCCGGAGCAGCGAAGTGGACGTCTCGGATCTCGGCTCGCGAAACTACAGTGCCCG 1612

Qy       1194 CACAGACTTCTACTGCCTGGTGTCCAAGGATGATATGTAG 1233
              |||||||||||||||||||||||||||||||||||||||
Db       1613 CACAGACTTCTACTGCCTGGTGTCCAAGGATGATATGTAG 1652


RESULT 13
ABK40277
ID   ABK40277 standard; cDNA; 1649 BP.
XX
AC   ABK40277;
XX
DT   15-JUL-2002  (first entry)
XX
DE   cDNA encoding human PRO342 polypeptide.
XX
```

KW Human; PRO; benign tumour; malignant tumour; lymphoid malignancy;
KW leukaemia; neuronal disorder; stromal disorder; blastocoelic disorder;
KW inflammatory disorder; immune disorder; angiogenic disorder;
KW gene therapy; cytostatic; neuroprotective; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200153486-A1.
XX
PD 26-JUL-2001.
XX
PF 11-FEB-2000; 2000WO-US003565.
XX
PR 08-MAR-1999; 99WO-US005028.
PR 11-MAR-1999; 99US-0123972P.
PR 11-MAY-1999; 99US-0133459P.
PR 02-JUN-1999; 99WO-US012252.
PR 22-JUN-1999; 99US-0140650P.
PR 22-JUN-1999; 99US-0140653P.
PR 20-JUL-1999; 99US-0144758P.
PR 26-JUL-1999; 99US-0145698P.
PR 28-JUL-1999; 99US-0146222P.
PR 17-AUG-1999; 99US-0149395P.
PR 31-AUG-1999; 99US-0151689P.
PR 01-SEP-1999; 99WO-US020111.
PR 15-SEP-1999; 99WO-US021090.
PR 30-NOV-1999; 99WO-US028313.
PR 01-DEC-1999; 99WO-US028301.
PR 01-DEC-1999; 99WO-US028634.
PR 05-JAN-2000; 2000WO-US000219.
XX
PA (GETH) GENENTECH INC.
XX
PI Ashkenazi AJ, Goddard A, Godowski PJ, Gurney AL, Hillan KJ;
PI Marsters SA, Pan J, Pitti RM, Roy MA, Smith V, Stone DM;
PI Watanabe CK, Wood WI;
XX
DR WPI; 2002-205567/26.
DR P-PSDB; AAU86151.
XX
PT Thirty five nucleic acids encoding PRO polypeptides, useful for treating
PT benign or malignant tumors, leukemias and lymphoid malignancies,
PT inflammatory, angiogenic and immunologic disorders.
XX
PS Claim 50; Fig 47; 302pp; English.
XX
CC The present invention relates to the isolation of novel human PRO
CC polypeptides and the polynucleotide sequences encoding them. The PRO
CC polypeptides, agonists, antagonists or anti-PRO antibodies are useful for
CC treating benign or malignant tumours (e.g. renal, kidney, bladder,
CC breast, etc), leukaemias and lymphoid malignancies, other disorders such
CC as neuronal, glial, astrocytal, hypothalamic, glandular, macrophagal,
CC stromal and blastocoelic disorders, inflammatory, immune and angiogenic
CC disorders. The polynucleotide sequences are also useful in gene therapy.
CC ABK40254-ABK40288 encode for the human PRO polypeptides of the invention
XX
SQ Sequence 1649 BP; 267 A; 554 C; 526 G; 302 T; 0 U; 0 Other;

Query Match 86.9%; Score 1071.8; DB 6; Length 1649;
Best Local Similarity 99.8%; Pred. No. 1.7e-204;
Matches 1073; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 ATGCCAGGTGTCTGTGATAGGGCCCCTGACTTCCTCTCCCCGTCTGAAGACCAGGTGCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        104 ATGCCAGGTGTCTGTGATAGGGCCCCTGACTTCCTCTCCCCGTCTGAAGACCAGGTGCTG 163

Qy         61 AGGCCTGCCTTGGGCAGCTCAGTGGCTCTGAACTGCACGGCTTGGGTAGTCTCTGGGCCC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        164 AGGCCTGCCTTGGGCAGCTCAGTGGCTCTGAACTGCACGGCTTGGGTAGTCTCTGGGCCC 223

Qy        121 CACTGCTCCCTGCCTTCAGTCCAGTGGCTGAAAGACGGGCTTCCATTGGGAATTGGGGGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        224 CACTGCTCCCTGCCTTCAGTCCAGTGGCTGAAAGACGGGCTTCCATTGGGAATTGGGGGC 283

Qy        181 CACTACAGCCTCCACGAGTACTCCTGGGTCAAGGCCAACCTGTCAGAGGTGCTTGTGTCC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        284 CACTACAGCCTCCACGAGTACTCCTGGGTCAAGGCCAACCTGTCAGAGGTGCTTGTGTCC 343

Qy        241 AGTGTCCTGGGGGTCAACGTGACCAGCACTGAAGTCTATGGGGCCTTCACCTGCTCCATC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        344 AGTGTCCTGGGGGTCAACGTGACCAGCACTGAAGTCTATGGGGCCTTCACCTGCTCCATC 403

Qy        301 CAGAACATCAGCTTCTCCTCCTTCACTCTTCAGAGAGCTGGCCCTACAAGCCACGTGGCT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        404 CAGAACATCAGCTTCTCCTCCTTCACTCTTCAGAGAGCTGGCCCTACAAGCCACGTGGCT 463

Qy        361 GCGGTGCTGGCCTCCCTCCTGGTCCTGCTGGCCCTGCTGCTGGCCGCCCTGCTCTATGTC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        464 GCGGTGCTGGCCTCCCTCCTGGTCCTGCTGGCCCTGCTGCTGGCCGCCCTGCTCTATGTC 523

Qy        421 AAGTGCCGTCTCAACGTGCTGCTCTGGTACCAGGACGCGTATGGGGAGGTGGAGATAAAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        524 AAGTGCCGTCTCAACGTGCTGCTCTGGTACCAGGACGCGTATGGGGAGGTGGAGATAAAC 583

Qy        481 GACGGGAAGCTCTACGACGCCTACGTCTCCTACAGCGACTGCCCCGAGGACCGCAAGTTC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        584 GACGGGAAGCTCTACGACGCCTACGTCTCCTACAGCGACTGCCCCGAGGACCGCAAGTTC 643

Qy        541 GTGAACTTCATCCTAAAGCCGCAGCTGGAGCGGCGTCGGGGCTACAAGCTCTTCCTGGAC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        644 GTGAACTTCATCCTAAAGCCGCAGCTGGAGCGGCGTCGGGGCTACAAGCTCTTCCTGGAC 703

Qy        601 GACCGCGACCTCCTGCCGCGCGCTGAGCCCTCCGCCGACCTCTTGGTGAACCTGAGCCGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        704 GACCGCGACCTCCTGCCGCGCGCTGAGCCCTCCGCCGACCTCTTGGTGAACCTGAGCCGC 763

Qy        661 TGCCGACGCCTCATCGTGGTGCTTTCGGACGCCTTCCTGAGCCGGGCCTGGTGCAGCCAC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        764 TGCCGACGCCTCATCGTGGTGCTTTCGGACGCCTTCCTGAGCCGGGCCTGGTGCAGCCAC 823

Qy        721 AGCTTCCGGGAGGGCCTGTGCCGGCTGCTGGAGCTCACCCGCAGACCCATCTTCATCACC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        824 AGCTTCCGGGAGGGCCTGTGCCGGCTGCTGGAGCTCACCCGCAGACCCATCTTCATCACC 883

Qy        781 TTCGAGGGCCAGAGGCGCGACCCCGCGCACCCGGCGCTCCGCCTGCTGCGCCAGCACCGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        884 TTCGAGGGCCAGAGGCGCGACCCCGCGCACCCGGCGCTCCGCCTGCTGCGCCAGCACCGC 943

Qy        841 CACCTGGTGACCTTGCTGCTCTGGAGGCCCGGCTCCGTGACTCCTTCCTCCGATTTTTGG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        944 CACCTGGTGACCTTGCTGCTCTGGAGGCCCGGCTCCGTGACTCCTTCCTCCGATTTTTGG 1003

Qy        901 AAAGAAGTGCAGCTGGCGCTGCCGCGGAAGGTGCGGTACAGGCCGGTGGAAGGAGACCCC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1004 AAAGAAGTGCAGCTGGCGCTGCCGCGGAAGGTGCGGTACAGGCCGGTGGAAGGAGACCCC 1063

Qy        961 CAGACGCAGCTGCAGGACGACAAGGACCCCATGCTGATTCTTCGAGGCCGAGTCCCTGAG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1064 CAGACGCAGCTGCAGGACGACAAGGACCCCATGCTGATTCTTCGAGGCCGAGTCCCTGAG 1123

Qy       1021 GGCCGGGCCCTGGACTCAGAGGTGGACCCGGACCCTGAGGGCGACCTGGGTGTCC 1075
              |||||||||||||||||||||||||||||||||||||||||||||||||| | |
Db       1124 GGCCGGGCCCTGGACTCAGAGGTGGACCCGGACCCTGAGGGCGACCTGGGTATGC 1178
```

RESULT 14
ADJ37328
ID ADJ37328 standard; cDNA; 1649 BP.
XX
AC ADJ37328;
XX
DT 22-APR-2004 (first entry)
XX
DE Human tumour therapy associated PRO342 cDNA.
XX
KW cytostatic; gene therapy; PRO; PRO197; PRO207; PRO226; PRO232; PRO243;

```
KW   PRO256; PRO269; PRO274; PRO304; PRO339; PRO1558; PRO779; PRO1185;
KW   PRO1245; PRO1759; PRO5775; PRO7133; PRO7168; PRO5725; PRO202; PRO206;
KW   PRO264; PRO313; PRO342; PRO542; PRO773; PRO861; PRO1216; PRO1686;
KW   PRO1800; PRO3562; PRO9850; PRO539; PRO4316; PRO4980; cancer; tumour;
KW   neoplastic cell growth; neoplastic cell proliferation; carcinoma;
KW   lymphoma; blastoma; sarcoma; leukaemia; gene; ss.
XX
OS   Homo sapiens.
XX
PN   US2003211096-A1.
XX
PD   13-NOV-2003.
XX
PF   02-AUG-2002; 2002US-00211858.
XX
PR   31-AUG-1999;   99US-0151689P.
PR   11-FEB-2000; 2000WO-US003565.
PR   09-AUG-2001; 2001US-00927796.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Ashkenazi AJ,  Goddard A,  Godowski PJ,  Gurney AL,  Hillan KJ;
PI   Marsters SA,  Pan J,  Pitti RM,  Roy MA,  Smith V,  Stone DM;
PI   Watanabe CK,  Wood WI;
XX
DR   WPI; 2003-901564/82.
DR   P-PSDB; ADJ37329.
XX
PT   New isolated PRO polypeptides, useful as targets for the diagnosis,
PT   prevention and treatment of cancers, e.g. lymphoma, blastoma, sarcoma or
PT   leukemia, and as predictors of the prognosis of tumor treatment.
XX
PS   Claim 50; SEQ ID NO 47; 307pp; English.
XX
CC   The invention describes an isolated PRO polypeptide. The PRO polypeptide:
CC   has at least 80% amino acid sequence identity to: (1) any one of 35 fully
CC   defined sequences of 104-954 amino acids (designated P1-P35) given in the
CC   specification, with or without its associated signal peptide; (2) an
CC   extracellular domain of any one of the polypeptides of P1-P35, with or
CC   without its associated signal peptide; or (3) an amino acid sequence
CC   encoded by the full-length coding sequence of the DNA deposited under
CC   ATCC accession number 209284, 209358, 203376, 209250, 209508, 209379,
CC   209397, 209786, 209482, 209490, 203312, 55820, 203096, 203155, 203465,
CC   PTA-255, PTA-618, PTA-545, PTA-256, 203538, 203661, 203835 or PTA-43; or
CC   scores at least 80% positives when compared to any one of the sequences
CC   of P1-P35. Specifically claimed are 35 PRO polypeptides, i.e. PRO197,
CC   PRO207, PRO226, PRO232, PRO243, PRO256, PRO269, PRO274, PRO304, PRO339,
CC   PRO1558, PRO779, PRO1185, PRO1245, PRO1759, PRO5775, PRO7133, PRO7168,
CC   PRO5725, PRO202, PRO206, PRO264, PRO313, PRO342, PRO542, PRO773, PRO861,
CC   PRO1216, PRO1686, PRO1800, PRO3562, PRO9850, PRO539, PRO4316 and PRO4980
CC   polypeptides. The PRO polypeptides are useful as targets for the
CC   diagnosis, prevention and treatment of cancers, and as predictors of the
CC   prognosis of tumour treatment. The nucleic acid molecules, antibodies and
CC   antagonists are useful for diagnosing and treating neoplastic cell growth
CC   and proliferation, e.g. carcinoma, lymphoma, blastoma, sarcoma or
CC   leukaemia. The antibodies may be used in screening assays for drug
CC   candidates. This sequence encodes a PRO protein useful in the treatment
CC   of cancers.
XX
SQ   Sequence 1649 BP; 267 A; 554 C; 526 G; 302 T; 0 U; 0 Other;

  Query Match             86.9%;  Score 1071.8;  DB 10;  Length 1649;
  Best Local Similarity   99.8%;  Pred. No. 1.7e-204;
  Matches 1073;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 ATGCCAGGTGTCTGTGATAGGGCCCCTGACTTCCTCTCCCCGTCTGAAGACCAGGTGCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        104 ATGCCAGGTGTCTGTGATAGGGCCCCTGACTTCCTCTCCCCGTCTGAAGACCAGGTGCTG 163

Qy         61 AGGCCTGCCTTGGGCAGCTCAGTGGCTCTGAACTGCACGGCTTGGGTAGTCTCTGGGCCC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        164 AGGCCTGCCTTGGGCAGCTCAGTGGCTCTGAACTGCACGGCTTGGGTAGTCTCTGGGCCC 223

Qy        121 CACTGCTCCCTGCCTTCAGTCCAGTGGCTGAAAGACGGGCTTCCATTGGGAATTGGGGGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        224 CACTGCTCCCTGCCTTCAGTCCAGTGGCTGAAAGACGGGCTTCCATTGGGAATTGGGGGC 283

Qy        181 CACTACAGCCTCCACGAGTACTCCTGGGTCAAGGCCAACCTGTCAGAGGTGCTTGTGTCC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        284 CACTACAGCCTCCACGAGTACTCCTGGGTCAAGGCCAACCTGTCAGAGGTGCTTGTGTCC 343

Qy        241 AGTGTCCTGGGGGTCAACGTGACCAGCACTGAAGTCTATGGGGCCTTCACCTGCTCCATC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        344 AGTGTCCTGGGGGTCAACGTGACCAGCACTGAAGTCTATGGGGCCTTCACCTGCTCCATC 403

Qy        301 CAGAACATCAGCTTCTCCTCCTTCACTCTTCAGAGAGCTGGCCCTACAAGCCACGTGGCT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        404 CAGAACATCAGCTTCTCCTCCTTCACTCTTCAGAGAGCTGGCCCTACAAGCCACGTGGCT 463

Qy        361 GCGGTGCTGGCCTCCCTCCTGGTCCTGCTGGCCCTGCTGCTGGCCGCCCTGCTCTATGTC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        464 GCGGTGCTGGCCTCCCTCCTGGTCCTGCTGGCCCTGCTGCTGGCCGCCCTGCTCTATGTC 523

Qy        421 AAGTGCCGTCTCAACGTGCTGCTCTGGTACCAGGACGCGTATGGGGAGGTGGAGATAAAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        524 AAGTGCCGTCTCAACGTGCTGCTCTGGTACCAGGACGCGTATGGGGAGGTGGAGATAAAC 583

Qy        481 GACGGGAAGCTCTACGACGCCTACGTCTCCTACAGCGACTGCCCCGAGGACCGCAAGTTC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        584 GACGGGAAGCTCTACGACGCCTACGTCTCCTACAGCGACTGCCCCGAGGACCGCAAGTTC 643

Qy        541 GTGAACTTCATCCTAAAGCCGCAGCTGGAGCGGCGTCGGGGCTACAAGCTCTTCCTGGAC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        644 GTGAACTTCATCCTAAAGCCGCAGCTGGAGCGGCGTCGGGGCTACAAGCTCTTCCTGGAC 703

Qy        601 GACCGCGACCTCCTGCCGCGCGCTGAGCCCTCCGCCGACCTCTTGGTGAACCTGAGCCGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        704 GACCGCGACCTCCTGCCGCGCGCTGAGCCCTCCGCCGACCTCTTGGTGAACCTGAGCCGC 763

Qy        661 TGCCGACGCCTCATCGTGGTGCTTTCGGACGCCTTCCTGAGCCGGGCCTGGTGCAGCCAC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        764 TGCCGACGCCTCATCGTGGTGCTTTCGGACGCCTTCCTGAGCCGGGCCTGGTGCAGCCAC 823

Qy        721 AGCTTCCGGGAGGGCCTGTGCCGGCTGCTGGAGCTCACCCGCAGACCCATCTTCATCACC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        824 AGCTTCCGGGAGGGCCTGTGCCGGCTGCTGGAGCTCACCCGCAGACCCATCTTCATCACC 883

Qy        781 TTCGAGGGCCAGAGGCGCGACCCCGCGCACCCGGCGCTCCGCCTGCTGCGCCAGCACCGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        884 TTCGAGGGCCAGAGGCGCGACCCCGCGCACCCGGCGCTCCGCCTGCTGCGCCAGCACCGC 943

Qy        841 CACCTGGTGACCTTGCTGCTCTGGAGGCCCGGCTCCGTGACTCCTTCCTCCGATTTTTGG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        944 CACCTGGTGACCTTGCTGCTCTGGAGGCCCGGCTCCGTGACTCCTTCCTCCGATTTTTGG 1003

Qy        901 AAAGAAGTGCAGCTGGCGCTGCCGCGGAAGGTGCGGTACAGGCCGGTGGAAGGAGACCCC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1004 AAAGAAGTGCAGCTGGCGCTGCCGCGGAAGGTGCGGTACAGGCCGGTGGAAGGAGACCCC 1063

Qy        961 CAGACGCAGCTGCAGGACGACAAGGACCCCATGCTGATTCTTCGAGGCCGAGTCCCTGAG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1064 CAGACGCAGCTGCAGGACGACAAGGACCCCATGCTGATTCTTCGAGGCCGAGTCCCTGAG 1123

Qy       1021 GGCCGGGCCCTGGACTCAGAGGTGGACCCGGACCCTGAGGGCGACCTGGGTGTCC 1075
              ||||||||||||||||||||||||||||||||||||||||||||||||||| | |
Db       1124 GGCCGGGCCCTGGACTCAGAGGTGGACCCGGACCCTGAGGGCGACCTGGGTATGC 1178

RESULT 15
ADG68252
ID   ADG68252 standard; cDNA; 1649 BP.
XX
AC   ADG68252;
XX
DT   11-MAR-2004  (first entry)
XX
```

DE Human PRO polypeptide cDNA #24.
XX
KW neoplastic tumour; lung; colon; breast; prostate; rectal; cervical;
KW liver; gene therapy; Human; ss; gene.
XX
OS Homo sapiens.
XX
PN US2003170228-A1.
XX
PD 11-SEP-2003.
XX
PF 02-AUG-2002; 2002US-00210951.
XX
PR 31-AUG-1999; 99US-0151689P.
PR 11-FEB-2000; 2000WO-US003565.
PR 09-AUG-2001; 2001US-00927796.
XX
PA (GETH) GENENTECH INC.
XX
PI Ashkenazi AJ, Goddard A, Godowski PJ, Gurney AL, Hillan KJ;
PI Marsters SA, Pan J, Pitti RM, Roy MA, Smith V, Stone DM;
PI Watanabe CK, Wood WI;
XX
DR WPI; 2004-020650/02.
DR P-PSDB; ADG68253.
XX
PT New isolated antibodies binding PRO polypeptides, useful for diagnosing,
PT prognosticating and/or treating neoplastic tumors, such as lung, colon,
PT breast, prostate, rectal, cervical and liver tumors.
XX
PS Claim 50; SEQ ID NO 47; 308pp; English.
XX
CC The invention relates to an isolated antibody that binds to a PRO
CC polypeptide. The methods and compositions of the present invention are
CC useful for diagnosing, prognosticating and/or treating neoplastic
CC tumours, such as lung, colon, breast, prostate, rectal, cervical and
CC liver tumours. The PRO polypeptides are also useful as molecular weight
CC markers, or for chromosome identification. The PRO genes are useful as
CC hybridisation probes, or for screening libraries of Human cDNA, genomic
CC DNA or mRNA. The PRO genes may also be used in gene therapy, particularly
CC for replacing a defective gene. The present sequence is used in the
CC exemplification of the invention.
XX
SQ Sequence 1649 BP; 267 A; 554 C; 526 G; 302 T; 0 U; 0 Other;

Query Match 86.9%; Score 1071.8; DB 12; Length 1649;
Best Local Similarity 99.8%; Pred. No. 1.7e-204;
Matches 1073; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 ATGCCAGGTGTCTGTGATAGGGCCCCTGACTTCCTCTCCCCGTCTGAAGACCAGGTGCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        104 ATGCCAGGTGTCTGTGATAGGGCCCCTGACTTCCTCTCCCCGTCTGAAGACCAGGTGCTG 163

Qy         61 AGGCCTGCCTTGGGCAGCTCAGTGGCTCTGAACTGCACGGCTTGGGTAGTCTCTGGGCCC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        164 AGGCCTGCCTTGGGCAGCTCAGTGGCTCTGAACTGCACGGCTTGGGTAGTCTCTGGGCCC 223

Qy        121 CACTGCTCCCTGCCTTCAGTCCAGTGGCTGAAAGACGGGCTTCCATTGGGAATTGGGGGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        224 CACTGCTCCCTGCCTTCAGTCCAGTGGCTGAAAGACGGGCTTCCATTGGGAATTGGGGGC 283

Qy        181 CACTACAGCCTCCACGAGTACTCCTGGGTCAAGGCCAACCTGTCAGAGGTGCTTGTGTCC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        284 CACTACAGCCTCCACGAGTACTCCTGGGTCAAGGCCAACCTGTCAGAGGTGCTTGTGTCC 343

Qy        241 AGTGTCCTGGGGGTCAACGTGACCAGCACTGAAGTCTATGGGGCCTTCACCTGCTCCATC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        344 AGTGTCCTGGGGGTCAACGTGACCAGCACTGAAGTCTATGGGGCCTTCACCTGCTCCATC 403

Qy        301 CAGAACATCAGCTTCTCCTCCTTCACTCTTCAGAGAGCTGGCCCTACAAGCCACGTGGCT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        404 CAGAACATCAGCTTCTCCTCCTTCACTCTTCAGAGAGCTGGCCCTACAAGCCACGTGGCT 463

Qy        361 GCGGTGCTGGCCTCCCTCCTGGTCCTGCTGGCCCTGCTGCTGGCCGCCCTGCTCTATGTC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        464 GCGGTGCTGGCCTCCCTCCTGGTCCTGCTGGCCCTGCTGCTGGCCGCCCTGCTCTATGTC 523

Qy        421 AAGTGCCGTCTCAACGTGCTGCTCTGGTACCAGGACGCGTATGGGGAGGTGGAGATAAAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        524 AAGTGCCGTCTCAACGTGCTGCTCTGGTACCAGGACGCGTATGGGGAGGTGGAGATAAAC 583

Qy        481 GACGGGAAGCTCTACGACGCCTACGTCTCCTACAGCGACTGCCCCGAGGACCGCAAGTTC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        584 GACGGGAAGCTCTACGACGCCTACGTCTCCTACAGCGACTGCCCCGAGGACCGCAAGTTC 643

Qy        541 GTGAACTTCATCCTAAAGCCGCAGCTGGAGCGGCGTCGGGGCTACAAGCTCTTCCTGGAC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        644 GTGAACTTCATCCTAAAGCCGCAGCTGGAGCGGCGTCGGGGCTACAAGCTCTTCCTGGAC 703

Qy        601 GACCGCGACCTCCTGCCGCGCGCTGAGCCCTCCGCCGACCTCTTGGTGAACCTGAGCCGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        704 GACCGCGACCTCCTGCCGCGCGCTGAGCCCTCCGCCGACCTCTTGGTGAACCTGAGCCGC 763

Qy        661 TGCCGACGCCTCATCGTGGTGCTTTCGGACGCCTTCCTGAGCCGGGCCTGGTGCAGCCAC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        764 TGCCGACGCCTCATCGTGGTGCTTTCGGACGCCTTCCTGAGCCGGGCCTGGTGCAGCCAC 823

Qy        721 AGCTTCCGGGAGGGCCTGTGCCGGCTGCTGGAGCTCACCCGCAGACCCATCTTCATCACC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        824 AGCTTCCGGGAGGGCCTGTGCCGGCTGCTGGAGCTCACCCGCAGACCCATCTTCATCACC 883

Qy        781 TTCGAGGGCCAGAGGCGCGACCCCGCGCACCCGGCGCTCCGCCTGCTGCGCCAGCACCGC 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        884 TTCGAGGGCCAGAGGCGCGACCCCGCGCACCCGGCGCTCCGCCTGCTGCGCCAGCACCGC 943

Qy        841 CACCTGGTGACCTTGCTGCTCTGGAGGCCCGGCTCCGTGACTCCTTCCTCCGATTTTTGG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        944 CACCTGGTGACCTTGCTGCTCTGGAGGCCCGGCTCCGTGACTCCTTCCTCCGATTTTTGG 1003

Qy        901 AAAGAAGTGCAGCTGGCGCTGCCGCGGAAGGTGCGGTACAGGCCGGTGGAAGGAGACCCC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1004 AAAGAAGTGCAGCTGGCGCTGCCGCGGAAGGTGCGGTACAGGCCGGTGGAAGGAGACCCC 1063

Qy        961 CAGACGCAGCTGCAGGACGACAAGGACCCCATGCTGATTCTTCGAGGCCGAGTCCCTGAG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1064 CAGACGCAGCTGCAGGACGACAAGGACCCCATGCTGATTCTTCGAGGCCGAGTCCCTGAG 1123

Qy       1021 GGCCGGGCCCTGGACTCAGAGGTGGACCCGGACCCTGAGGGCGACCTGGGTGTCC 1075
              ||||||||||||||||||||||||||||||||||||||||||||||||||||| | |
Db       1124 GGCCGGGCCCTGGACTCAGAGGTGGACCCGGACCCTGAGGGCGACCTGGGTATGC 1178
```

Search completed: July 29, 2005, 22:29:56
Job time : 763 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: July 29, 2005, 21:44:31 ; Search time 4597 Seconds
(without alignments)
10209.538 Million cell updates/sec

Title: US-09-598-443-1
Perfect score: 1233
Sequence: 1 atgccaggtgtctgtgatag.....tgtccaaggatgatatgtag 1233

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_htc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	806.4	65.4	1043	5	BM915345	BM915345 AGENCOURT
2	733	59.4	763	4	BI517905	BI517905 603041985
3	725.2	58.8	941	4	BG740350	BG740350 602634105
c 4	708.2	57.4	983	5	BX400205	BX400205 BX400205
5	681.4	55.3	821	2	BE797163	BE797163 601587171
6	660.4	53.6	852	4	BI521062	BI521062 603081652
7	640.6	52.0	839	5	BX369506	BX369506 BX369506
8	639.2	51.8	749	4	BI821947	BI821947 603039758
9	616	50.0	808	4	BI769076	BI769076 603057053
10	615.2	49.9	621	7	CV026019	CV026019 3834 Full
c 11	606.6	49.2	791	4	BI518182	BI518182 603041985
12	601.4	48.8	1106	5	BM924346	BM924346 AGENCOURT
13	589.8	47.8	845	2	BE792803	BE792803 601584677
14	585.4	47.5	1003	5	BX341382	BX341382 BX341382
15	575	46.6	584	4	BM714596	BM714596 UI-E-EJ0-
16	572.6	46.4	774	5	BP446530	BP446530 BP446530
17	567.2	46.0	590	7	CN309086	CN309086 170004182
18	556.8	45.2	941	4	BI518544	BI518544 603061669
19	556.2	45.1	744	5	BP446681	BP446681 BP446681
20	556.2	45.1	930	5	BX400206	BX400206 BX400206
21	546.8	44.3	1067	4	BI752545	BI752545 603021956
22	541	43.9	983	4	BI759065	BI759065 603042860
23	535.8	43.5	796	4	BG742077	BG742077 602633540
24	534	43.3	704	7	CF146731	CF146731 UI-HF-CB0
25	522.4	42.4	801	4	BI770099	BI770099 603053464
c 26	516.2	41.9	720	4	BM675360	BM675360 UI-E-EJ0-
c 27	516.2	41.9	722	5	BM982596	BM982596 UI-CF-EN1
28	515.4	41.8	700	5	BP448975	BP448975 BP448975
29	513	41.6	783	4	BI820133	BI820133 603037143
30	504.2	40.9	730	5	BX926048	BX926048 BX926048
31	496.4	40.3	821	4	BI758455	BI758455 603022612
32	488.4	39.6	513	2	BF239133	BF239133 601905529
33	488.4	39.6	851	5	BX382402	BX382402 BX382402
34	486.8	39.5	1116	5	BQ053327	BQ053327 AGENCOURT
35	486	39.4	925	4	BI905958	BI905958 603062841
36	482.2	39.1	970	2	BE250059	BE250059 600943075
37	476.4	38.6	659	6	CD673345	CD673345 fg23h01.y
c 38	474.4	38.5	786	5	BX341381	BX341381 BX341381
39	471.6	38.2	1131	5	BM922484	BM922484 AGENCOURT
40	468.8	38.0	822	4	BI909088	BI909088 603070070
41	467.8	37.9	892	2	BE612529	BE612529 601452037
42	466.6	37.8	868	5	BQ232590	BQ232590 AGENCOURT
c 43	465.8	37.8	794	5	BX352106	BX352106 BX352106
44	454.8	36.9	458	1	AV655410	AV655410 AV655410
45	453	36.7	616	4	BM715423	BM715423 UI-E-EJ0-

ALIGNMENTS

RESULT 1
BM915345
LOCUS BM915345 1043 bp mRNA linear EST 12-MAR-2002
DEFINITION AGENCOURT_6701721 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5481517 5', mRNA sequence.
ACCESSION BM915345
VERSION BM915345.1 GI:19365724
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1043)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTP
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2006 row: b column: 14
High quality sequence stop: 611.
FEATURES Location/Qualifiers
source 1. .1043
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5481517"
/tissue_type="amelanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_41"
/note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."
ORIGIN

Query Match 65.4%; Score 806.4; DB 5; Length 1043;

```
  Best Local Similarity   92.5%;  Pred. No. 8.4e-165;
  Matches  896;  Conservative    0;  Mismatches   32;  Indels   41;  Gaps    3;

Qy        148 CTGAAAGACGGGCTTCCATTGGGAATTGGGGGCCACTACAGCCTCCACGAGTACTCCTGG 207
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CTGAAAGACGGGCTTCCATTGGGAATTGGGGGCCACTACAGCCTCCACGAGTACTCCTGG 60

Qy        208 GTCAAGGCCAACCTGTCAGAGGTGCTTGTGTCCAGTGTCCTGGGGGTCAACGTGACCAGC 267
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTCAAGGCCAACCTGTCAGAGGTGCTTGTGTCCAGTGTCCTGGGGGTCAACGTGACCAGC 120

Qy        268 ACTGAAGTCTATGGGGCCTTCACCTGCTCCATCCAGAACATCAGCTTCTCCTCCTTCACT 327
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACTGAAGTCTATGGGGCCTTCACCTGCTCCATCCAGAACATCAGCTTCTCCTCCTTCACT 180

Qy        328 CTTCAGAGAGCTGGCCCTACAAGCCACGTGGCTGCGGTGCTGGCCTCCCTCCTGGTCCTG 387
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTTCAGAGAGCTGGCCCTACAAGCCACGTGGCTGCGGTGCTGGCCTCCCTCCTGGTCCTG 240

Qy        388 CTGGCCCTGCTGCTGGCCGCCCTGCTCTATGTCAAGTGCCGTCTCAACGTGCTGCTCTGG 447
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTGGCCCTGCTGCTGGCCGCCCTGCTCTATGTCAAGTGCCGTCTCAACGTGCTGCTCTGG 300

Qy        448 TACCAGGACGCGTATGGGGAGGTGGAGATAAACGACGGGAAGCTCTACGACGCCTACGTC 507
              ||||||||||||||||||||||||||||||||
Db        301 TACCAGGACGCGTATGGGGAGGTGGAGATAAAC-------------------------- 333

Qy        508 TCCTACAGCGACTGCCCCGAGGACCGCAAGTTCGTGAACTTCATCCTAAAGCCGCAGCTG 567
                      ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        334 -------GCGACTGCCCCGAGGACCGCAAGTTCGTGAACTTCATCCTAAAGCCGCAGCTG 386

Qy        568 GAGCGGCGTCGGGGCTACAAGCTCTTCCTGGACGACCGCGACCTCCTGCCGCGCGCTGAG 627
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        387 GAGCGGCGTCGGGGCTACAAGCTCTTCCTGGACGACCGCGACCTCCTGCCGCGCGCTGAG 446

Qy        628 CCCTCCGCCGACCTCTTGGTGAACCTGAGCCGCTGCCGACGCCTCATCGTGGTGCTTTCG 687
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        447 CCCTCCGCCGACCTCTTGGTGAACCTGAGCCGCTGCCGACGCCTCATCGTGGTGCTTTCG 506

Qy        688 GACGCCTTCCTGAGCCGGGCCTGGTGCAGCCACAGCTTCCGGGAGGGCCTGTGCCGGCTG 747
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        507 GACGCCTTCCTGAGCCGGGCCTGGTGCAGCCACAGCTTCCGGGAGGGCCTGTGCCGGCTG 566

Qy        748 CTGGAGCTCACCCGCAGACCCATCTTCATCACCTTCGAGGGCCAGAGGCGCGACCCCGCG 807
              |||||||||||||||||| |||||||||||||||| ||||||||||||||||||||||||
Db        567 CTGGAGCTCACCCGCAGACCTATCTTCATCACCTTCGAAGGCCAGAGGCGCGACCCCGCG 626

Qy        808 CACCCGGCGCTCCGCCTGCTGCGCCAGCACCGCCACCTGGTGACCTTGCTGCTCTGGAGG 867
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        627 CACCCGGCGCTCCGCCTGCTGCGCCAGCACCGCCACCTGGTGACCTTGCTGCTCTGGAGG 686

Qy        868 CCCGGCTCCGTGACTCCTTCCTCCGATTTTTGGAAAGAAGTGCAGCTGGCGCTGCCGCGG 927
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        687 CCCGGCTCCGTGACTCCTTCCTCCGATTTTTGGAAAG-AGTGCAGCTGGCGCTGCCGCGG 745

Qy        928 AAGGTGCGGTACAGGCCGGTGGAAGGAGACCCCCAGACGCAGCTGCAGGACGACAAGGAC 987
              |||||||||||||||||||| | ||||||||||||||||||||||||||||||||||||
Db        746 GAGGTGCGGTACAGGCCGGTGGGAAGAGACCCCCAGACGCAGCTGCAGGACGACAAGGAC 805

Qy        988 CCCATGCTGATTCTTCGAGGCCGAGTCCCTGAGGGCCGGGCCCTGGACTCAGAGGTGGAC 1047
              |||||||||||||||||||  |||||||||||| || ||  | ||||||| ||||||
Db        806 CCCATGCTGATTCTTCGAGGCGAAGTCCCTGAGGGCGGGCCCTGGNACTCAGAAGTGGAC 865

Qy       1048 CCGGACCCTGAGGGCGACCTGGGTGTCCGGGGG------CCTGTTTTTGGAGAGCCATCA 1101
              || ||||    |||| ||  ||| |  | ||||          ||||||||||||||||
Db        866 CCCGACCTGAGGGGCCACTGGGGGGGGCCGGGGGGGCCCGGGTTTTTTGGAGAGCCATCA 925

Qy       1102 GCTCCACCG 1110
              ||| |  ||
Db        926 GCTTCCACG 934


RESULT 2
```

```
BI517905
LOCUS       BI517905                763 bp    mRNA    linear   EST 29-AUG-2001
DEFINITION  603041985F1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5182556 5',
            mRNA sequence.
ACCESSION   BI517905
VERSION     BI517905.1  GI:15342697
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 763)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11455  row: i  column: 21
            High quality sequence stop: 739.
FEATURES             Location/Qualifiers
     source          1. .763
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5182556"
                     /lab_host="DH10B"
                     /clone_lib="NIH_MGC_116"
                     /note="Organ: pooled colon, kidney, stomach; Vector:
                     pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
                     source anonymous pool of 3 colons, age 26 yo male, 49 yo
                     female, 71 yo male colon; 46 yo male kidney, and pool of 2
                     stomachs, 62 yo male and 70 yo female.  Library is
                     oligo-dT primed and  directionally cloned (EcoRV site is
                     destroyed upon cloning).  Average insert size 1.4 kb,
                     insert size range 1-3 kb.  Library is normalized and
                     enriched for full-length clones and was constructed by C.
                     Gruber  (Invitrogen).  Research Genetics tracking code
                     023. Note: this is a NIH_MGC Library."
ORIGIN

  Query Match            59.4%;  Score 733;  DB 4;  Length 763;
  Best Local Similarity  99.6%;  Pred. No. 7.5e-149;
  Matches  755;  Conservative    0;  Mismatches    1;  Indels    2;  Gaps    2;

Qy        272 AAGTCTATGGGGCCTTCACCTGCTCCATCCAGAACATCAGCTTCTCCTCCTTCACTCTTC 331
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AAGTCTATGGGGCCTTCACCTGCTCCATCCAGAACATCAGCTTCTCCTCCTTCACTCTTC 60

Qy        332 AGAGAGCTGGCCCTACAAGCCACGTGGCTGCGGTGCTGGCCTCCCTCCTGGTCCTGCTGG 391
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGAGAGCTGGCCCTACAAGCCACGTGGCTGCGGTGCTGGCCTCCCTCCTGGTCCTGCTGG 120

Qy        392 CCCTGCTGCTGGCCGCCCTGCTCTATGTCAAGTGCCGTCTCAACGTGCTGCTCTGGTACC 451
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCCTGCTGCTGGCCGCCCTGCTCTATGTCAAGTGCCGTCTCAACGTGCTGCTCTGGTACC 180

Qy        452 AGGACGCGTATGGGGAGGTGGAGATAAACGACGGGAAGCTCTACGACGCCTACGTCTCCT 511
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AGGACGCGTATGGGGAGGTGGAGATAAACGACGGGAAGCTCTACGACGCCTACGTCTCCT 240

Qy        512 ACAGCGACTGCCCCGAGGACCGCAAGTTCGTGAACTTCATCCTAAAGCCGCAGCTGGAGC 571
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ACAGCGACTGCCCCGAGGACCGCAAGTTCGTGAACTTCATCCTAAAGCCGCAGCTGGAGC 300

Qy        572 GGCGTCGGGGCTACAAGCTCTTCCTGGACGACCGCGACCTCCTGCCGCGCGCTGAGCCCT 631
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        301 GGCGTCGGGGCTACAAGCTCTTCCTGGACGACCGCGACCTCCTGCCGCGCGCTGAGCCCT 360

Qy        632 CCGCCGACCTCTTGGTGAACCTGAGCCGCTGCCGACGCCTCATCGTGGTGCTTTCGGACG 691
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCGCCGACCTCTTGGTGAACCTGAGCCGCTGCCGACGCCTCATCGTGGTGCTTTCGGACG 420

Qy        692 CCTTCCTGAGCCGGGCCTGGTGCAGCCACAGCTTCCGGGAGGGCCTGTGCCGGCTGCTGG 751
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CCTTCCTGAGCCGGGCCTGGTGCAGCCACAGCTTCCGGGAGGGCCTGTGCCGGCTGCTGG 480

Qy        752 AGCTCACCCGCAGACCCATCTTCATCACCTTCGAGGGCCAGAGGCGCGACCCCGCGCACC 811
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AGCTCACCCGCAGACCCATCTTCATCACCTTCGAGGGCCAGAGGCGCGACCCCGCGCACC 540

Qy        812 CGGCGCTCCGCCTGCTGCGCCAGCACCGCCACCTGGTGACCTTGCTGCTCTGGAGGCCCG 871
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 CGGCGCTCCGCCTGCTGCGCCAGCACCGCCACCTGGTGACCTTGCTGCTCTGGAGGCCCG 600

Qy        872 GCTCCGTGACTCCTTCCTCCGATTTTTGGAAAGAAGTGCAGCTGGCGCTGCCGCGGAAGG 931
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        601 GCTCCGTGACTCCTTCCTCCGA-TTTTGGAAAGAAGTGCAGCTGGCGCTGCCGCGGAAGG 659

Qy        932 TGCGGTACAGGCCGGTGGAAGGAGACCCCCAGACGCAGCTGCAGGACGACAAGGACCCCA 991
              |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        660 TGCGGTACAGGCCGGTGGAAGGAGACCCCCAGACGCAG-TGCAGGACGACAAGGACCCCA 718

Qy        992 TGCTGATTCTTCGAGGCCGAGTCCCTGAGGGCCGGGCC 1029
              |||||||||||||||||||||| ||||||||||||||
Db        719 TGCTGATTCTTCGAGGCCGAGTCNCTGAGGGCCGGGCC 756
```

RESULT 3
BG740350

```
LOCUS       BG740350                941 bp    mRNA    linear   EST 15-MAY-2001
DEFINITION  602634105F1 NCI_CGAP_Skn3 Homo sapiens cDNA clone IMAGE:4779337 5',
            mRNA sequence.
ACCESSION   BG740350
VERSION     BG740350.1  GI:14051003
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 941)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: James Cleaver, M.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA
            Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10636  row: i  column: 02
            High quality sequence stop: 856.
FEATURES             Location/Qualifiers
     source          1. .941
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4779337"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /clone_lib="NCI_CGAP_Skn3"
                     /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.5kb.  Library constructed by Life
                     Technologies. Note: this is a NCI_CGAP Library."
ORIGIN
```

Query Match 58.8%; Score 725.2; DB 4; Length 941;
Best Local Similarity 98.1%; Pred. No. 3.8e-147;
Matches 808; Conservative 0; Mismatches 8; Indels 8; Gaps 7;

```
Qy        411 GCTCTATGTCAAGTGCCGTCTCAACGTGCTGCTCTGGTACCAGGACGCGTATGGGGAGGT 470
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GCTCTATGTCAAGTGCCGTCTCAACGTGCTGCTCTGGTACCAGGACGCGTATGGGGAGGT 60

Qy        471 GGAGATAAACGACGGGAAGCTCTACGACGCCTACGTCTCCTACAGCGACTGCCCCGAGGA 530
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GGAGATAAACGACGGGAAGCTCTACGACGCCTACGTCTCCTACAGCGACTGCCCCGAGGA 120

Qy        531 CCGCAAGTTCGTGAACTTCATCCTAAAGCCGCAGCTGGAGCGGCGTCGGGGCTACAAGCT 590
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCGCAAGTTCGTGAACTTCATCCTAAAGCCGCAGCTGGAGCGGCGTCGGGGCTACAAGCT 180

Qy        591 CTTCCTGGACGACCGCGACCTCCTGCCGCGCGCTGAGCCCTCCGCCGACCTCTTGGTGAA 650
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTTCCTGGACGACCGCGACCTCCTGCCGCGCGCTGAGCCCTCCGCCGACCTCTTGGTGAA 240

Qy        651 CCTGAGCCGCTGCCGACGCCTCATCGTGGTGCTTTCGGACGCCTTCCTGAGCCGGGCCTG 710
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCTGAGCCGCTGCCGACGCCTCATCGTGGTGCTTTCGGACGCCTTCCTGAGCCGGGCCTG 300

Qy        711 GTGCAGCCACAGCTTCCGGGAGGGCCTGTGCCGGCTGCTGGAGCTCACCCGCAGACCCAT 770
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GTGCAGCCACAGCTTCCGGGAGGGCCTGTGCCGGCTGCTGGAGCTCACCCGCAGACCCAT 360

Qy        771 CTTCATCACCTTCGAGGGCCAGAGGCGCGACCCCGCGCACCCGGCGCTCCGCCTGCTGCG 830
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CTTCATCACCTTCGAGGGCCAGAGGCGCGACCCCGCGCACCCGGCGCTCCGCCTGCTGCG 420

Qy        831 CCAGCACCGCCACCTGGTGACCTTGCTGCTCTGGAGGCCCGGCTCCGTGACTCCTTCCTC 890
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CCAGCACCGCCACCTGGTGACCTTGCTGCTCTGGAGGCCCGGCTCCGTGACTCCTTCCTC 480

Qy        891 CGATTTTTGGAAAGAAGTGCAGCTGGCGCTGCCGCGGAAGGTGCGGTACAGGCCGGTGG- 949
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db        481 CGATTTTTGGAAAGAAGTGCAGCTGGCGCTGCCGCGGAAGGTGCGGTACAGGCCGGTGGC 540

Qy        950 AAGGAGACCCCCAGACGCAGCTGC-AGGACGACAAGGACCCCATGCTGATTCTTCGAGGC 1008
              |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db        541 AAGGAGACCCCCAGACGCAGCTGCAAGGACGACAAGGACCCCATGCTGATTCTTCGAGGC 600

Qy       1009 CGAGTCCCTGAGGGCCGGGCCCTGG-ACTCAGAGGTGGACCCGGACCCTGAGGGCGACCT 1067
              |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db        601 CGAGTCCCTGAGGGCCGGGCCCTGGAACTCAGAGGTGGACCCGGACCCTGAGGGCGACCT 660

Qy       1068 GGGTGTCCGGGGGCCTGTTTTTGGAGAGCCATCAGCTCCACCGCACACCAGTGGGGTCTC 1127
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GGGTGTCCGGGGGCCTGTTTTTGGAGAGCCATCAGCTCCACCGCACACCAGTGGGGTCTC 720

Qy       1128 GCTGGG-AGAGAGCCGGAGCA--GCGAAGTGGACGTCTCGGATCTCGGCTCGCGAAACTA 1184
              |||||| |||||||||||| ||     ||||||||||||||||||||||||||||||||||
Db        721 GCTGGGAAGAGAGCCGGAACAAGCGAAAGTGGACGTCTCGGATCTCGGCTCGCGAAACTA 780

Qy       1185 CAGTGCCCGCACA-GACTTCTACTGCCT-GGTGTCCAAGGATGA 1226
              |||||| |||||| |  |||||||||||| |||||||| |||||
Db        781 CAGTGCGCGCACAGGTTTTCTACTGCCTGGGTGTCCAACGATGA 824
```

RESULT 4
BX400205/c

```
LOCUS       BX400205                983 bp    mRNA    linear   EST 28-APR-2004
DEFINITION  BX400205 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
            Homo sapiens cDNA clone CS0DJ005YE04 3-PRIME, mRNA sequence.
ACCESSION   BX400205
VERSION     BX400205.2  GI:46847557
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 983)
```

```
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On May 13, 2003 this sequence version replaced gi:30618108.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
FEATURES             Location/Qualifiers
     source          1. .983
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DJ005YE04"
                     /cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
                     /cell_line="JURKAT"
                     /clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT
                     10-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN

  Query Match             57.4%;  Score 708.2;  DB 5;  Length 983;
  Best Local Similarity   96.8%;  Pred. No. 1.9e-143;
  Matches  729;  Conservative   11;  Mismatches   11;  Indels    2;  Gaps    2;

Qy        481 GACGGGAAGCTCTACGACGCCTACGTCTCCTACAGCGACTGCCCCGAGGACCGCAAGTTC 540
              ||||||||||||:|||::|| ||| |: ||::|:||| |||||||||||||||:||||||||
Db        890 GACGGGAAGCTYTACKRCGKCTATGKTTCYKAYAGCTACTGCCCCGAGGACYGCAAGTTC 831

Qy        541 GTGAACTTCATCCTAAAGCCGCAGCTGGAGCGGCGTCGGGGCTACAAGCTCTTCCTGGAC 600
              ||||||||||||:|||||||||| ||||||||||||||||||||||||||||||||||||
Db        830 GTGAACTTCATCYTAAAGCCGCAGTTGGAGCGGCGTCGGGGCTACAAGCTCTTCCTGGAC 771

Qy        601 GACCGCGACCTCCTGCCGCGCGCTGAGCCCTCCGCCGACCTCTTGGTGAACCTGAGCCGC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        770 GACCGCGACCTCCTGCCGCGCGCTGAGCCCTCCGCCGACCTCTTGGTGAACCTGAGCCGC 711

Qy        661 TGCCGACGCCTCATCGTGGTGCTTTCGGACGCCTTCCTGAGCCGGGCCTGGTGCAGCCAC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        710 TGCCGACGCCTCATCGTGGTGCTTTCGGACGCCTTCCTGAGCCGGGCCTGGTGCAGCCAC 651

Qy        721 AGCTTCCGGGAGGGCCTGTGCCGGCTGCTGGAGCTCACCCGCAGACCCATCTTCATCACC 780
              |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        650 AGCTTCCGGGAGGGCCTGTGCCGNCTGCTGGAGCTCACCCGCAGACCCATCTTCATCACC 591

Qy        781 TTCGAGGGCCAGAGGCGCGACCCCGCGCACCCGGCGCTCCGCCTGCTGCGCCAGCACCGC 840
              |||||||||||||||||||||||||||||||||||||||||| :|||||||||||||||
Db        590 TTCGAGGGCCAGAGGCGCGACCCCGCGCACCCGGCGCTCCGCCT-STGCGCCAGCACCGC 532

Qy        841 CACCTGGTGACCTTGCTGCTCTGGAGGCCCGGCTCCGTGACTCCTTCCTCCGATTTTTGG 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        531 CACCTGGTGACCTTGCTGCTCTGGAGGCCCGGCTCCGTGACTCCTTCCTCCGATTTTTGG 472

Qy        901 AAAGAAGTGCAGCTGGCGCTGCCGCGGAAGGTGCGGTACAGGCCGGTGGAAGGAGACCCC 960
              |||||||  |||||||| |||||||||||||||| |||||||| ||||||||||||||||
Db        471 AAAGAAGNTCAGCTGGC-CTGCCGCGGAAGGTGCAGTACAGGCCTGTGGAAGGAGACCCC 413

Qy        961 CAGACGCAGCTGCAGGACGACAAGGACCCCATGCTGATTCTTCGAGGCCGAGTCCCTGAG 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        412 CAGACGCAGCTGCAGGACGACAAGGACCCCATGCTGATTCTTCGAGGCCGAGTCCCTGAG 353

Qy       1021 GGCCGGGCCCTGGACTCAGAGGTGGACCCGGACCCTGAGGGCGACCTGGGTGTCCGGGGG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        352 GGCCGGGCCCTGGACTCAGAGGTGGACCCGGACCCTGAGGGCGACCTGGGTGTCCGGGGG 293

Qy       1081 CCTGTTTTTGGAGAGCCATCAGCTCCACCGCACACCAGTGGGGTCTCGCTGGGAGAGAGC 1140
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        292 CCTGTCTTTGGAGAGCCATCAGCTCCACCGCACACCAGTGGGGTCTCGCTGGGAGAGAGC 233

Qy       1141 CGGAGCAGCGAAGTGGACGTCTCGGATCTCGGCTCGCGAAACTACAGTGCCCGCACAGAC 1200
              ||||||||||||||||||||||||||||||:|||||||||||||||||||||||||||||
Db        232 CGGAGCAGCGAAGTGGACGTCTCGGATCTCRGCTCGCGAAACTACAGTGCCCGCACAGAC 173

Qy       1201 TTCTACTGCCTGGTGTCCAAGGATGATATGTAG 1233
              |||||||||||||||||||||||||||||||||
Db        172 TTCTACTGCCTGGTGTCCAAGGATGATATGTAG 140


RESULT 5
BE797163
LOCUS       BE797163                821 bp    mRNA    linear   EST 20-SEP-2000
DEFINITION  601587171F1 NIH_MGC_7 Homo sapiens cDNA clone IMAGE:3941561 5',
            mRNA sequence.
ACCESSION   BE797163
VERSION     BE797163.1  GI:10218361
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 821)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
            Plate: LLCM793  row: m  column: 18
            High quality sequence stop: 793.
FEATURES             Location/Qualifiers
     source          1. .821
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3941561"
                     /tissue_type="small cell carcinoma"
                     /cell_line="MGC3"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_7"
                     /note="Organ: lung; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming.  Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G).  Size-selected >500bp for average
                     insert size 1.8kb.  Library constructed by Ling Hong in
                     the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using  ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies)."
ORIGIN

  Query Match             55.3%;  Score 681.4;  DB 2;  Length 821;
  Best Local Similarity   94.4%;  Pred. No. 1.3e-137;
  Matches  728;  Conservative    0;  Mismatches   41;  Indels    2;  Gaps    2;

Qy        206 GGGTCAAGGCCAACCTGTCAGAGGTGCTTGTGTCCA-GTGTCCTGGGGGTCAACGTGACC 264
              |||||||||||||| |||||||||||||||||||| |||||||||||||||||||||||
Db         14 GGGTCAAGGCCAA-CTGTCAGAGGTGCTTGTGTCCATGTGTCCTGGGGGTCAACGTGACC 72

Qy        265 AGCACTGAAGTCTATGGGGCCTTCACCTGCTCCATCCAGAACATCAGCTTCTCCTCCTTC 324
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         73 AGCACTGAAGTCTATGGGGCCTTCACCTGCTCCATCCAGAACATCAGCTTCTCCTCCTTC 132

Qy        325 ACTCTTCAGAGAGCTGGCCCTACAAGCCACGTGGCTGCGGTGCTGGCCTCCCTCCTGGTC 384
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        133 ACTCTTCAGAGAGCTGGCCCTACAAGCCACGTGGCTGCGGTGCTGGCCTCCCTCCTGGTC 192

Qy        385 CTGCTGGCCCTGCTGCTGGCCGCCCTGCTCTATGTCAAGTGCCGTCTCAACGTGCTGCTC 444
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        193 CTGCTGGCCCTGCTGCTGGCCGCCCTGCTCTATGTCAAGTGCCGTCTCAACGTGCTGCTC 252

Qy        445 TGGTACCAGGACGCGTATGGGGAGGTGGAGATAAACGACGGGAAGCTCTACGACGCCTAC 504
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        253 TGGTACCAGGACGCGTATGGGGAGGTGGAGATAAACGACGGGAAGCTCTACGACGCCTAC 312

Qy        505 GTCTCCTACAGCGACTGCCCCGAGGACCGCAAGTTCGTGAACTTCATCCTAAAGCCGCAG 564
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        313 GTCTCCTACAGCGACTGCCCCGAGGACCGCAAGTTCGTGAACTTCATCCTAAAGCCGCAG 372

Qy        565 CTGGAGCGGCGTCGGGGCTACAAGCTCTTCCTGGACGACCGCGACCTCCTGCCGCGCGCT 624
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        373 CTGGAGCGGCGTCGGGGCTACAAGCTCTTCCTGGACGACCGCGACCTCCTGCCGCGCGCT 432

Qy        625 GAGCCCTCCGCCGACCTCTTGGTGAACCTGAGCCGCTGCCGACGCCTCATCGTGGTGCTT 684
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        433 GAGCCCTCCGCCGACCTCTTGGTGAACCTGAGCCGCTGCCGACGCCTCATCGTGGTGCTT 492

Qy        685 TCGGACGCCTTCCTGAGCCGGGCCTGGTGCAGCCACAGCTTCCGGGAGGGCCTGTGCCGG 744
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        493 TCGGACGCCTTCCTGAGCCGGGCCTGGTGCAGCCACAGCTTCCGGGAGGGCCTGTGCCGG 552

Qy        745 CTGCTGGAGCTCACCCGCAGACCCATCTTCATCACCTTCGAGGGCCAGAGGCGCGACCCC 804
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db        553 CTGCTGGAGCTCACCCGCAGACCCATCTTCATCACCTTCGAGGGCCAGAGGCGCGAACCC 612

Qy        805 GCGCACCCGGCGCTCCGCCTGCTGCGCCAGCACCGCCACCTGGTGACCTTGCTGCTCTGG 864
              ||||||| |||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        613 GCGCACCTGGCGCTCCGCCTGCTGCGCCAGCACCGGCACCTGGTGACCTTGCTGCTCTGG 672

Qy        865 AGGCCCGGCTCCGTGACTCCTTCCTCCGATTTTTGGAAAGAAGTGCAGCTGGCGCTGCCG 924
              ||| |||||||     | ||||||||||||||||||||||||||| |||||||||
Db        673 AGGGCCGGCTCGTGACTCCTTTCCTCCGATTTTTGGAAAGAAGTGCAGGTGGCGCTGCGC 732

Qy        925 CGGAAGGTGCGGTACAGGCCGGTGGAAGGAGACCCCCAGACGCAGCTGCAG 975
              | ||  ||   |      | | | |  ||||| ||| |||  |||||
Db        733 GGAAGTGCGGAACAGGCCGGTGGGAAGGAGAACCCCAAGAGGCAATTGCAG 783
```

RESULT 6
BI521062
LOCUS BI521062 852 bp mRNA linear EST 29-AUG-2001
DEFINITION 603081652F1 NIH_MGC_120 Homo sapiens cDNA clone IMAGE:5220901 5', mRNA sequence.
ACCESSION BI521062
VERSION BI521062.1 GI:15345854
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 852)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11555 row: g column: 14
High quality sequence start: 3
High quality sequence stop: 780.
FEATURES Location/Qualifiers
source 1. .852
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5220901"
/lab_host="DH10B"
/clone_lib="NIH_MGC_120"
/note="Organ: pooled pancreas and spleen; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of spleen and pancreas from 28 yo male. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.5 kb, insert size range 1-2.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 025. Note: this is a NIH_MGC Library."
ORIGIN

Query Match 53.6%; Score 660.4; DB 4; Length 852;
Best Local Similarity 98.7%; Pred. No. 4.7e-133;
Matches 697; Conservative 0; Mismatches 6; Indels 3; Gaps 3;

```
Qy          1 ATGCCAGGTGTCTGTGATAGGGCCCCTGACTTCCTCTCCCCGTCTGAAGACCAGGTGCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        148 ATGCCAGGTGTCTGTGATAGGGCCCCTGACTTCCTCTCCCCGTCTGAAGACCAGGTGCTG 207

Qy         61 AGGCCTGCCTTGGGCAGCTCAGTGGCTCTGAACTGCACGGCTTGGGTAGTCTCTGGGCCC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        208 AGGCCTGCCTTGGGCAGCTCAGTGGCTCTGAACTGCACGGCTTGGGTAGTCTCTGGGCCC 267

Qy        121 CACTGCTCCCTGCCTTCAGTCCAGTGGCTGAAAGACGGGCTTCCATTGGGAATTGGGGGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        268 CACTGCTCCCTGCCTTCAGTCCAGTGGCTGAAAGACGGGCTTCCATTGGGAATTGGGGGC 327

Qy        181 CACTACAGCCTCCACGAGTACTCCTGGGTCAAGGCCAACCTGTCAGAGGTGCTTGTGTCC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        328 CACTACAGCCTCCACGAGTACTCCTGGGTCAAGGCCAACCTGTCAGAGGTGCTTGTGTCC 387

Qy        241 AGTGTCCTGGGGGTCAACGTGACCAGCACTGAAGTCTATGGGGCCTTCACCTGCTCCATC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        388 AGTGTCCTGGGGGTCAACGTGACCAGCACTGAAGTCTATGGGGCCTTCACCTGCTCCATC 447

Qy        301 CAGAACATCAGCTTCTCCTCCTTCACTCTTCAGAGAGCTGGCCCTACAAGCCACGTGGCT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        448 CAGAACATCAGCTTCTCCTCCTTCACTCTTCAGAGAGCTGGCCCTACAAGCCACGTGGCT 507

Qy        361 GCGGTGCTGGCCTCCCTCCTGGTCCTGCTGGCCCTGCTGCTGGCCGCCCTGCTCTATGTC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        508 GCGGTGCTGGCCTCCCTCCTGGTCCTGCTGGCCCTGCTGCTGGCCGCCCTGCTCTATGTC 567

Qy        421 AAGTGCCGTCTCAACGTGCTGCTCTGGTACCAGGACGCGTATGGGGAGGTGGAGATAAAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        568 AAGTGCCGTCTCAACGTGCTGCTCTGGTACCAGGACGCGTATGGGGAGGTGGAGATAAAC 627

Qy        481 GACGGGAAGCTCTACGACGCCTACGTCTCCTACAGCGACTGCCCCGAGGACCGCAAGTTC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        628 GACGGGAAGCTCTACGACGCCTACGTCTCCTACAGCGACTGCCCCGAGGACCGCAAGTTC 687

Qy        541 GTGAACTTCATCCTAAAGCCGCAGCTGGA-GCGGCGTCGGGGCTACAAGCTCTTCCTGGA 599
              |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db        688 GTGAACTTCATCCTAAAGCCGCAGCTGGACGCGGCGTCGGGGCTACAAGCTCTTCCTGGA 747

Qy        600 CGACCGCGACCTCCTGCCGCGCGCTGAGCCCTCCGCCGACCTCTTGGTGAACCTGA-GCC 658
              |||||| |||||||||||||||||||||||||||||| |||||||||||||||||| |||
Db        748 CGACCG-GACCTCCTGCCGCGCGCTGAGCCCTCCGGCGACCTCTTGGTGAACCTGAGGCC 806

Qy        659 GCTGCCGACGCCTCATCGTGGTGCTTTCGGACGCCTTCCTGAGCCG 704
              | |||||||||||||||||||||||| ||||||||| | | |||||
Db        807 GGTGCCGACGCCTCATCGTGGTGCTTCCGGACGCCTCCTTAAGCCG 852
```

RESULT 7
BX369506
LOCUS BX369506 839 bp mRNA linear EST 28-APR-2004

```
DEFINITION  BX369506 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
            Homo sapiens cDNA clone CS0DJ015YD03 5-PRIME, mRNA sequence.
ACCESSION   BX369506
VERSION     BX369506.2  GI:46833010
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 839)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On May 8, 2003 this sequence version replaced gi:30457818.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
FEATURES             Location/Qualifiers
     source          1. .839
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DJ015YD03"
                     /cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
                     /cell_line="JURKAT"
                     /clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT
                     10-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN

  Query Match            52.0%;  Score 640.6;  DB 5;  Length 839;
  Best Local Similarity  99.4%;  Pred. No. 9.5e-129;
  Matches  643;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy        587 AGCTCTTCCTGGACGACCGCGACCTCCTGCCGCGCGCTGAGCCCTCCGCCGACCTCTTGG 646
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         12 AGCTCTTCCTGGACGACCGCGACCTCCTGCCGCGCGCTGAGCCCTCCGCCGACCTCTTGG 71

Qy        647 TGAACCTGAGCCGCTGCCGACGCCTCATCGTGGTGCTTTCGGACGCCTTCCTGAGCCGGG 706
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         72 TGAACCTGAGCCGCTGCCGACGCCTCATCGTGGTGCTTTCGGACGCCTTCCTGAGCCGGG 131

Qy        707 CCTGGTGCAGCCACAGCTTCCGGGAGGGCCTGTGCCGGCTGCTGGAGCTCACCCGCAGAC 766
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        132 CCTGGTGCAGCCACAGCTTCCGGGAGGGCCTGTGCCGGCTGCTGGAGCTCACCCGCAGAC 191

Qy        767 CCATCTTCATCACCTTCGAGGGCCAGAGGCGCGACCCCGCGCACCCGGCGCTCCGCCTGC 826
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        192 CCATCTTCATCACCTTCGAGGGCCAGAGGCGCGACCCCGCGCACCCGGCGCTCCGCCTGC 251

Qy        827 TGCGCCAGCACCGCCACCTGGTGACCTTGCTGCTCTGGAGGCCCGGCTCCGTGACTCCTT 886
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        252 TGCGCCAGCACCGCCACCTGGTGACCTTGCTGCTCTGGAGGCCCGGCTCCGTGACTCCTT 311

Qy        887 CCTCCGATTTTTGGAAAGAAGTGCAGCTGGCGCTGCCGCGGAAGGTGCGGTACAGGCCGG 946
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||| |||
Db        312 CCTCCGATTTTTGGAAAGAAGTGCAGCTGGCGCTGCCGCGGAAGGTGCAGTACAGGCCTG 371

Qy        947 TGGAAGGAGACCCCCAGACGCAGCTGCAGGACGACAAGGACCCCATGCTGATTCTTCGAG 1006
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        372 TGGAAGGAGACCCCCAGACGCAGCTGCAGGACGACAAGGACCCCATGCTGATTCTTCGAG 431

Qy       1007 GCCGAGTCCCTGAGGGCCGGGCCCTGGACTCAGAGGTGGACCCGGACCCTGAGGGCGACC 1066
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        432 GCCGAGTCCCTGAGGGCCGGGCCCTGGACTCAGAGGTGGACCCGGACCCTGAGGGCGACC 491

Qy       1067 TGGGTGTCCGGGGGCCTGTTTTTGGAGAGCCATCAGCTCCACCGCACACCAGTGGGGTCT 1126
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        492 TGGGTGTCCGGGGGCCTGTCTTTGGAGAGCCATCAGCTCCACCGCACACCAGTGGGGTCT 551

Qy       1127 CGCTGGGAGAGAGCCGGAGCAGCGAAGTGGACGTCTCGGATCTCGGCTCGCGAAACTACA 1186
              |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        552 CGCTGGAAGAGAGCCGGAGCAGCGAAGTGGACGTCTCGGATCTCGGCTCGCGAAACTACA 611

Qy       1187 GTGCCCGCACAGACTTCTACTGCCTGGTGTCCAAGGATGATATGTAG 1233
              |||||||||||||||||||||||||||||||||||||||||||||
Db        612 GTGCCCGCACAGACTTCTACTGCCTGGTGTCCAAGGATGATATGTAG 658


RESULT 8
BI821947
LOCUS       BI821947                749 bp    mRNA    linear   EST 04-OCT-2001
DEFINITION  603039758F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5180511 5',
            mRNA sequence.
ACCESSION   BI821947
VERSION     BI821947.1  GI:15933497
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 749)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11450  row: d  column: 16
            High quality sequence stop: 739.
FEATURES             Location/Qualifiers
     source          1. .749
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5180511"
                     /lab_host="DH10B"
                     /clone_lib="NIH_MGC_115"
                     /note="Organ: pooled brain, lung, testis; Vector:
                     pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
                     source anonymous pool of 6 male brains, age range 23-27; 1
                     male lung, age 27; and 1 male testis, age 69.  Library is
                     oligo-dT primed and directionally cloned (EcoRV site is
                     destroyed upon  cloning).  Average insert size 1.8 kb,
                     insert size range 1-3 kb. Library is normalized and
                     enriched for full-length clones and was  constructed by C.
                     Gruber (Invitrogen).  Research Genetics tracking  code
                     021. Note: this is a NIH_MGC Library."
ORIGIN

  Query Match            51.8%;  Score 639.2;  DB 4;  Length 749;
  Best Local Similarity  99.5%;  Pred. No. 1.9e-128;
  Matches  641;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy        203 CCTGGGTCAAGGCCAACCTGTCAGAGGTGCTTGTGTCCAGTGTCCTGGGGGTCAACGTGA 262
              ||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        106 CCATGGTCAAGGCCAACCTGTCAGAGGTGCTTGTGTCCAGTGTCCTGGGGGTCAACGTGA 165

Qy        263 CCAGCACTGAAGTCTATGGGGCCTTCACCTGCTCCATCCAGAACATCAGCTTCTCCTCCT 322
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        166 CCAGCACTGAAGTCTATGGGGCCTTCACCTGCTCCATCCAGAACATCAGCTTCTCCTCCT 225
```

```
Qy        323 TCACTCTTCAGAGAGCTGGCCCTACAAGCCACGTGGCTGCGGTGCTGGCCTCCCTCCTGG 382
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        226 TCACTCTTCAGAGAGCTGGCCCTACAAGCCACGTGGCTGCGGTGCTGGCCTCCCTCCTGG 285

Qy        383 TCCTGCTGGCCCTGCTGCTGGCCGCCCTGCTCTATGTCAAGTGCCGTCTCAACGTGCTGC 442
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        286 TCCTGCTGGCCCTGCTGCTGGCCGCCCTGCTCTATGTCAAGTGCCGTCTCAACGTGCTGC 345

Qy        443 TCTGGTACCAGGACGCGTATGGGGAGGTGGAGATAAACGACGGGAAGCTCTACGACGCCT 502
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        346 TCTGGTACCAGGACGCGTATGGGGAGGTGGAGATAAACGACGGGAAGCTCTACGACGCCT 405

Qy        503 ACGTCTCCTACAGCGACTGCCCCGAGGACCGCAAGTTCGTGAACTTCATCCTAAAGCCGC 562
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        406 ACGTCTCCTACAGCGACTGCCCCGAGGACCGCAAGTTCGTGAACTTCATCCTAAAGCCGC 465

Qy        563 AGCTGGAGCGGCGTCGGGGCTACAAGCTCTTCCTGGACGACCGCGACCTCCTGCCGCGCG 622
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        466 AGCTGGAGCGGCGTCGGGGCTACAAGCTCTTCCTGGACGACCGCGACCTCCTGCCGCGCG 525

Qy        623 CTGAGCCCTCCGCCGACCTCTTGGTGAACCTGAGCCGCTGCCGACGCCTCATCGTGGTGC 682
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        526 CTGAGCCCTCCGCCGACCTCTTGGTGAACCTGAGCCGCTGCCGACGCCTCATCGTGGTGC 585

Qy        683 TTTCGGACGCCTTCCTGAGCCGGGCCTGGTGCAGCCACAGCTTCCGGGAGGGCCTGTGCC 742
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        586 TTTCGGACGCCTTCCTGAGCCGGGCCTGGTGCAGCCACAGCTTCCGGGAGGGCCTGTGCC 645

Qy        743 GGCTGCTGGAGCTCACCCGCAGACCCATCTTCATCACCTTCGAGGGCCAGAGGCGCGACC 802
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db        646 GGCTGCTGGAGCTCACCCGCAGACCTATCTTCATCACCTTCGAGGGCCAGAGGCGCGACC 705

Qy        803 CCGCGCACCCGGCGCTCCGCCTGCTGCGCCAGCACCGCCACCTG 846
              |||||||||||||||||||||||||||||||||||||||||||||
Db        706 CCGCGCACCCGGCGCTCCGCCTGCTGCGCCAGCACCGCCACCTG 749
```

RESULT 9
BI769076
LOCUS BI769076 808 bp mRNA linear EST 25-SEP-2001
DEFINITION 603057053F1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5206380 5', mRNA sequence.
ACCESSION BI769076
VERSION BI769076.1 GI:15760654
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 808)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11517 row: j column: 13
High quality sequence stop: 807.
FEATURES Location/Qualifiers
source 1. .808
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5206380"
/lab_host="DH10B"
/clone_lib="NIH_MGC_122"
/note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of 24 week female lung, 16 week female spleen, and 20-22 week male spleens. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 026. Note: this is a NIH_MGC Library."
ORIGIN

Query Match 50.0%; Score 616; DB 4; Length 808;
Best Local Similarity 97.4%; Pred. No. 2.1e-123;
Matches 637; Conservative 0; Mismatches 15; Indels 2; Gaps 1;

```
Qy          4 CCAGGTGTCTGTGATAGGGCCCCTGACTTCCTCTCCCCGTCTGAAGACCAGGTGCTGAGG 63
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        155 CCAGGTGTCTGTGATAGGGCCCCTGACTTCCTCTCCCCGTCTGAAGACCAGGTGCTGAGG 214

Qy         64 CCTGCCTTGGGCAGCTCAGTGGCTCTGAACTGCACGGCTTGGGTAGTCTCTGGGCCCCAC 123
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        215 CCTGCCTTGGGCAGCTCAGTGGCTCTGAACTGCACGGCTTGGGTAGTCTCTGGGCCCCAC 274

Qy        124 TGCTCCCTGCCTTCAGTCCAGTGGCTGAAAGACGGGCTTCCATTGGGAATTGGGGGCCAC 183
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        275 TGCTCCCTGCCTTCAGTCCAGTGGCTGAAAGACGGGCTTCCATTGGGAATTGGGGGCCAC 334

Qy        184 TACAGCCTCCACGAGTACTCCTGGGTCAAGGCCAACCTGTCAGAGGTGCTTGTGTCCAGT 243
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        335 TACAGCCTCCACGAGTACTCCTGGGTCAAGGCCAACCTGTCAGAGGTGCTTGTGTCCAGT 394

Qy        244 GTCCTGGGGGTCAACGTGACCAGCACTGAAGTCTATGGGGCCTTCACCTGCTCCATCCAG 303
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        395 GTCCTGGGGGTCAACGTGACCAGCACTGAAGTCTATGGGGCCTTCACCTGCTCCATCCAG 454

Qy        304 AACATCAGCTTCTCCTCCTTCACTCTTCAGAGAGCTGGCCCTACAAGCCACGTGGCTGCG 363
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        455 AACATCAGCTTCTCCTCCTTCACTCTTCAGAGAGCTGGCCCTACAAGCCACGTGGCTGCG 514

Qy        364 GTGCTGGCCTCCCTCCTGGTCCTGCTGGCCCTGCTGCTGGCCGCCCTGCTCTATGTCAAG 423
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        515 GTGCTGGCCTCCCTCCTGGTCCTGCTGGCCCTGCTGCTGGCCGCCCTGCTCTATGTCAAG 574

Qy        424 TGCCGTCTCAACGTGCTGCTCTGGTACCAGGACGCGTATGGGGAGGTGGAGATAAACGAC 483
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        575 TGCCGTCTCAACGTGCTGCTCTGGTACCAGGACGCGTATGGGGAGGTGGAGATAAACGAC 634

Qy        484 GGGAAGCTCTACGACGCCTACGTCTCCTACAGCGACTGCCCCGAGGACCGCAAGTTCGTG 543
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        635 GGGAAGCTCTACGACGCCTACGTCTCCTACAGCGACTGCCCCGAGGACCGCAAGTTCGTG 694

Qy        544 AACTTCATCCTAAAGCCGCAGCTGGAGCGGCGTCGGGGCTACAA--GCTCTTCCTGGACG 601
              |||||||||||| |||||||||||||||||||||||||||||||       | | |||||||
Db        695 AACTTCATCCTACAGCCGCAGCTGGAGCGGCGTCGGGGCTACAAAGCTCTTCCTTGGACG 754

Qy        602 ACCGCGACCTCCTGCCGCGCGCTGAGCCCTCCGCCGACCTCTTGGTGAACCTGA 655
              |||| || |||||||||||||||||||||||||| ||||    | | |||||||||
Db        755 ACCGGGAACTCCTGCCGCGCGCTGAGCCCTCCGACGACTCTTGGTTGAACCTGA 808
```

RESULT 10
CV026019
LOCUS CV026019 621 bp mRNA linear EST 20-AUG-2004
DEFINITION 3834 Full Length cDNA from the Mammalian Gene Collection Homo sapiens cDNA 5' similar to BC003591, mRNA sequence.
ACCESSION CV026019
VERSION CV026019.1 GI:51483978
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```
REFERENCE   1  (bases 1 to 621)
  AUTHORS   Rual,J.F., Hirozane-Kishikawa,T., Hao,T., Bertin,N., Li,S.,
            Dricot,A., Li,N., Rosenberg,J., Lamesch,P., Vidalain,P.O.,
            Clingingsmith,T.R., Hartley,J.L., Esposito,D., Cheo,D., Moore,T.,
            Simmons,B., Sequerra,R., Bosak,S., Doucette-Stamm,L., Le Peuch,C.,
            Vandenhaute,J., Cusick,M.E., Albala,J.S., Hill,D.E. and Vidal,M.
  TITLE     Human ORFeome Version 1.1: a Platform for Reverse Proteomics
  JOURNAL   Genome Res. (2004) In press
COMMENT     Contact: Vidal M
            Marc Vidal Laboratory
            Dana Farber Cancer Institute
            1 Jimmy Fund Way Smith 858, BOSTON, MA 02115, USA
            Tel: 617 632 5180
            Fax: 617 632 5739
            Email: Marc_Vidal@dfci.harvard.edu
            ORF Sequence Tag (OST) of Gateway Entry construct. Each cloned ORF
            results from a PCR reaction using an MGC full-length cDNA as
            template DNA and ORF specific primers
            PCR PRimers
            FORWARD: ATGCCAGGTGTCTGTGAT
            BACKWARD: TACATATCATCCTTGGACACCAG
            Insert Length: 621   Std Error: 36.00
            Plate: 11010  row: 08  column: C
            Seq primer: ACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAAC
            High quality sequence start: 96
            High quality sequence stop: 620
            POLYA=No.
FEATURES              Location/Qualifiers
     source           1. .621
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /tissue_type="mixed"
                      /clone_lib="Full Length cDNA from the Mammalian Gene
                      Collection"
                      /note="Vector: mixed; The ORFs were PCR amplified from the
                      MGC (Mammalian Gene Collection) as of April 2004 and
                      cloned by recombinational Gateway cloning into pDONR223
                      Donor vector. Reference : MGC (Mammalian Gene Collection)
                      Program Team, Generation and Initial Analysis of more than
                      15,000 Full-Length Human and Mouse cDNA Sequences. PNAS,
                      2002, 99(26), 16899-16903"
ORIGIN

  Query Match             49.9%;  Score 615.2;  DB 7;  Length 621;
  Best Local Similarity   99.4%;  Pred. No. 3.1e-123;
  Matches  617;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 ATGCCAGGTGTCTGTGATAGGGCCCCTGACTTCCTCTCCCCGTCTGAAGACCAGGTGCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCCAGGTGTCTGTGATAGGGCCCCTGACTTCCTCTCCCCGTCTGAAGACCAGGTGCTG 60

Qy         61 AGGCCTGCCTTGGGCAGCTCAGTGGCTCTGAACTGCACGGCTTGGGTAGTCTCTGGGCCC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGGCCTGCCTTGGGCAGCTCAGTGGCTCTGAACTGCACGGCTTGGGTAGTCTCTGGGCCC 120

Qy        121 CACTGCTCCCTGCCTTCAGTCCAGTGGCTGAAAGACGGGCTTCCATTGGGAATTGGGGGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CACTGCTCCCTGCCTTCAGTCCAGTGGCTGAAAGACGGGCTTCCATTGGGAATTGGGGGC 180

Qy        181 CACTACAGCCTCCACGAGTACTCCTGGGTCAAGGCCAACCTGTCAGAGGTGCTTGTGTCC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CACTACAGCCTCCACGAGTACTCCTGGGTCAAGGCCAACCTGTCAGAGGTGCTTGTGTCC 240

Qy        241 AGTGTCCTGGGGGTCAACGTGACCAGCACTGAAGTCTATGGGGCCTTCACCTGCTCCATC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGTGTCCTGGGGGTCAACGTGACCAGCACTGAAGTCTATGGGGCCTTCACCTGCTCCATC 300

Qy        301 CAGAACATCAGCTTCTCCTCCTTCACTCTTCAGAGAGCTGGCCCTACAAGCCACGTGGCT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CAGAACATCAGCTTCTCCTCCTTCACTCTTCAGAGAGCTGGCCCTACAAGCCACGTGGCT 360

Qy        361 GCGGTGCTGGCCTCCCTCCTGGTCCTGCTGGCCCTGCTGCTGGCCGCCCTGCTCTATGTC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCGGTGCTGGCCTCCCTCCTGGTCCTGCTGGCCCTGCTGCTGGCCGCCCTGCTCTATGTC 420

Qy        421 AAGTGCCGTCTCAACGTGCTGCTCTGGTACCAGGACGCGTATGGGGAGGTGGAGATAAAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AAGTGCCGTCTCAACGTGCTGCTCTGGTACCAGGACGCGTATGGGGAGGTGGAGATAAAC 480

Qy        481 GACGGGAAGCTCTACGACGCCTACGTCTCCTACAGCGACTGCCCCGAGGACCGCAAGTTC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GACGGGAAGCTCTACGACGCCTACGTCTCCTACAGCGACTGCCCCGAGGACCGCAAGTTC 540

Qy        541 GTGAACTTCATCCTAAAGCCGCAGCTGGAGCGGCGTCGGGGCTACAAGCTCTTCCTGGAC 600
              |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db        541 GTGAACTTCATCCTANAGCCGCAGCTGGAGCGGCGTCGGGGCTACAAGCTCTTCCTGGAC 600

Qy        601 GACCGCGACCTCCTGCCGCGC 621
              |||||||||||||   |||||
Db        601 GACCGCGACCTCCCTGCGCGC 621


RESULT 11
BI518182/c
LOCUS       BI518182                791 bp    mRNA    linear   EST 29-AUG-2001
DEFINITION  603041985T1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5182556 3',
            mRNA sequence.
ACCESSION   BI518182
VERSION     BI518182.1  GI:15342974
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 791)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11455  row: i  column: 21
            High quality sequence start: 40
            High quality sequence stop: 733.
FEATURES              Location/Qualifiers
     source           1. .791
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:5182556"
                      /lab_host="DH10B"
                      /clone_lib="NIH_MGC_116"
                      /note="Organ: pooled colon, kidney, stomach; Vector:
                      pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
                      source anonymous pool of 3 colons, age 26 yo male, 49 yo
                      female, 71 yo male colon; 46 yo male kidney, and pool of 2
                      stomachs, 62 yo male and 70 yo female.  Library is
                      oligo-dT primed and  directionally cloned (EcoRV site is
                      destroyed upon cloning).  Average insert size 1.4 kb,
                      insert size range 1-3 kb.  Library is normalized and
                      enriched for full-length clones and was constructed by C.
                      Gruber  (Invitrogen).  Research Genetics tracking code
                      023. Note: this is a NIH_MGC Library."
ORIGIN

  Query Match             49.2%;  Score 606.6;  DB 4;  Length 791;
  Best Local Similarity   98.0%;  Pred. No. 2.3e-121;
  Matches  646;  Conservative    0;  Mismatches    9;  Indels    4;  Gaps    3;
```

```
Qy        579 GGGCTACAAGCTCTTCCTGGACGACCGCGACCTCCTGCCGCGCGCTGAGCCCTC-CGCCG 637
              ||||||||||||||||||||||| ||||||||||||| |||||||||||||||| |||||
Db        790 GGGCTACAAGCTCTTCCTGGACGTCCGCGACCTCCTGGCGCGCGCTGAGCCCTCTCGCCG 731

Qy        638 ACCTCTTGGTGAACCTGAGCCGCTGCCGACGCCTCATCGTGGTGCTTTCGGACGCCTTCC 697
              |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db        730 ACCTCTTGGTGAACCTGATCCGCTGCCGACGCCTCATCGTGGTGCTTTCGGACGCCTTCC 671

Qy        698 TGAGCCGGGCCTGGTGCAGCCACAGCTTCCGGGAGGGCCTGTGCCGGCTGCTGGAGCTCA 757
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        670 TGAGCCGGGCCTGGTGCAGCCACAGCTTCCGGGAGGGCCTGTGCCGGCTGCTGGAGCTCA 611

Qy        758 CCCGCAGACCCATCTTCATCACCTTCGAGGGCCAGAGGCGC-GACCCCGCGCACCCGGCG 816
              |||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db        610 CCCGCAGACCCATCTTCATCACCTTCGAGGGCCAGAGGCGCTGACCCCGCGCACCCGGCG 551

Qy        817 CTCCGCCT--GCTGCGCCAGCACCGCCACCTGGTGACCTTGCTGCTCTGGAGGCCCGGCT 874
              ||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        550 CTCCGCCTTGCTGCGGCCAGCACCGCCACCTGGTGACCTTGCTGCTCTGGAGGCCCGGCT 491

Qy        875 CCGTGACTCCTTCCTCCGATTTTTGGAAAGAAGTGCAGCTGGCGCTGCCGCGGAAGGTGC 934
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        490 CCGTGACTCCTTCCTCCGATTTTTGGAAAGAAGTGCAGCTGGCGCTGCCGCGGAAGGTGC 431

Qy        935 GGTACAGGCCGGTGGAAGGAGACCCCCAGACGCAGCTGCAGGACGACAAGGACCCCATGC 994
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        430 GGTACAGGCCGGTGGAAGGAGACCCCCAGACGCAGCTGCAGGACGACAAGGACCCCATGC 371

Qy        995 TGATTCTTCGAGGCCGAGTCCCTGAGGGCCGGGCCCTGGACTCAGAGGTGGACCCGGACC 1054
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db        370 TGATTCTTCGAGGCCGAGTCCCTGAGGGCCGGGCCCTGGACTCAGAGGTGGACCCGGAAC 311

Qy       1055 CTGAGGGCGACCTGGGTGTCCGGGGGCCTGTTTTTGGAGAGCCATCAGCTCCACCGCACA 1114
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        310 CTGAGGGCGACCTGGGTGTCCGGGGGCCTGTTTTTGGAGAGCCATCAGCTCCACCGCACA 251

Qy       1115 CCAGTGGGGTCTCGCTGGGAGAGAGCCGGAGCAGCGAAGTGGACGTCTCGGATCTCGGCT 1174
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 CCAGTGGGGTCTCGCTGGGAGAGAGCCGGAGCAGCGAAGTGGACGTCTCGGATCTCGGCT 191

Qy       1175 CGCGAAACTACAGTGCCCGCACAGACTTCTACTGCCTGGTGTCCAAGGATGATATGTAG 1233
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         190 CGCGAAACTACAGTGCCCGCACAGACTTCTACTGCCTGGTGTCCAAGGATGATATGTAG 132
```

RESULT 12
BM924346

```
LOCUS       BM924346               1106 bp    mRNA    linear   EST 12-MAR-2002
DEFINITION  AGENCOURT_6630582 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5760639
            5', mRNA sequence.
ACCESSION   BM924346
VERSION     BM924346.1  GI:19374725
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1106)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12807  row: p  column: 16
            High quality sequence stop: 655.
FEATURES             Location/Qualifiers
     source          1. .1106
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5760639"
                     /lab_host="DH10B"
                     /clone_lib="NIH_MGC_116"
                     /note="Organ: pooled colon, kidney, stomach; Vector:
                     pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
                     source anonymous pool of 3 colons, age 26 yo male, 49 yo
                     female, 71 yo male colon; 46 yo male kidney, and pool of 2
                     stomachs, 62 yo male and 70 yo female.  Library is
                     oligo-dT primed and  directionally cloned (EcoRV site is
                     destroyed upon cloning).  Average insert size 1.4 kb,
                     insert size range 1-3 kb.  Library is normalized and
                     enriched for full-length clones and was constructed by C.
                     Gruber  (Invitrogen).  Research Genetics tracking code
                     023. Note: this is a NIH_MGC Library."
ORIGIN
```

Query Match 48.8%; Score 601.4; DB 5; Length 1106;
Best Local Similarity 93.2%; Pred. No. 3.2e-120;
Matches 728; Conservative 0; Mismatches 41; Indels 12; Gaps 9;

```
Qy          1 ATGCCAGGTGTCTGTGATAGGGCCCCTGACTTCCTCTCCCCGTCTGAAGACCAGGTGCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        198 ATGCCAGGTGTCTGTGATAGGGCCCCTGACTTCCTCTCCCCGTCTGAAGACCAGGTGCTG 257

Qy         61 AGGCCTGCCTTGGGCAGCTCAGTGGCTCTGAACTGCACGGCTTGGGTAGTCTCTGGGCCC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        258 AGGCCTGCCTTGGGCAGCTCAGTGGCTCTGAACTGCACGGCTTGGGTAGTCTCTGGGCCC 317

Qy        121 CACTGCTCCCTGCCTTCAGTCCAGTGGCTGAAAGACGGGCTTCCATTGGGAATTGGGGGC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        318 CACTGCTCCCTGCCTTCAGTCCAGTGGCTGAAAGACGGGCTTCCATTGGGAATTGGGGGC 377

Qy        181 CACTACAGCCTCCACGAGTACTCCTGGGTCAAGGCCAACCTGTCAGAGGTGCTTGTGTCC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        378 CACTACAGCCTCCACGAGTACTCCTGGGTCAAGGCCAACCTGTCAGAGGTGCTTGTGTCC 437

Qy        241 AGTGTCCTGGGGGTCAACGTGACCAGCACTGAAGTCTATGGGGCCTTCACCTGCTCCATC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        438 AGTGTCCTGGGGGTCAACGTGACCAGCACTGAAGTCTATGGGGCCTTCACCTGCTCCATC 497

Qy        301 CAGAACATCAGCTTCTCCTCCTTCACTCTTCAGAGAGCTGGCCCTACAAGCCACGTGGCT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        498 CAGAACATCAGCTTCTCCTCCTTCACTCTTCAGAGAGCTGGCCCTACAAGCCACGTGGCT 557

Qy        361 GCGGTGCTGGCCTCCCTCCTGGTCCTGCTGGCCCTGCTGCTGGCCGCCCTGCTCTATGTC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        558 GCGGTGCTGGCCTCCCTCCTGGTCCTGCTGGCCCTGCTGCTGGCCGCCCTGCTCTATGTC 617

Qy        421 AAGTGCCGTCTCAACGTGCTGCTCTGGTACCAGGA-CGCGTATGGGGAGGTGGAGATAAA 479
              |||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db        618 AAGTGCCGTCTCAACGTGCTGCTCTGGTACCAGGACCGCGTATGGGGAGGTGGAGATAAA 677

Qy        480 CGACGGGAAGCTCTACGACGCCTACGTCTCCTACAGCGACTGCCCCGAGGACCGCAAGTT 539
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        678 CGACGGGAAGCTCTACGACGCCTACGTCTCCTACAGCGACTGCCCCGAGGACCGCAAGTT 737

Qy        540 CGTGAACTTCATCCTAAAGCCGCAGCTGGAGC-GGCGTCGGGGCTACAA--GCTCTTCCT 596
              |||||||||||||||||||||||||||||||| ||||||||||||||      |||||
Db        738 CGTGAACTTCATCCTAAAGCCGCAGCTGGAGCGGGCGTCGGGGCTACAAAGCTCTTTCCT 797

Qy        597 GGACGACC-GCGACCTCCTGCCGCGCGCTGAGCCCTCCGCCGACCTCTTGGTG--AACCT 653
              |||||||| |||| |||||||||||||||||||||||| | |||||||||  |||
Db        798 GGACGACCGGCGAACTCCTGCCGCGCGCTGAGCCCTCCGCCAAACTCTTGGTGGAAACTG 857

Qy        654 GAGCCGCTGCCGACGCCTCATCGTGGTG-CTTTCGGACGCCTT--CCTGAGCCGGGCCTG 710
              | |    |||||| |||||||||| |  |||||||||||||  |  ||||||||||||
Db        858 AACCGCTGGCCGACCCCTCATCGTGGGGTTTTTCGGACGCCTTTCCTGGAGCCGGGCCTG 917

Qy        711 GT-GCAGCCACA-GCTTCCGGGAGGGCCTGTGCCGGCTGCTGGAGCTCACCCGCAGACCC 768
```

```
Db        918 GTGGCAGCCACAGGCTTCCCGGAAGGGCCCTGGGCCCGGCTGGCGGGAAACTCAACCCGC 977

Qy        769 A 769
              |
Db        978 A 978


RESULT 13
BE792803
LOCUS       BE792803                845 bp    mRNA    linear   EST 20-SEP-2000
DEFINITION  601584677F1 NIH_MGC_7 Homo sapiens cDNA clone IMAGE:3939162 5',
            mRNA sequence.
ACCESSION   BE792803
VERSION     BE792803.1  GI:10214001
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 845)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
            Plate: LLCM787  row: i  column: 19
            High quality sequence stop: 790.
FEATURES             Location/Qualifiers
     source          1. .845
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3939162"
                     /tissue_type="small cell carcinoma"
                     /cell_line="MGC3"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_7"
                     /note="Organ: lung; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming.  Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G).  Size-selected >500bp for average
                     insert size 1.8kb.  Library constructed by Ling Hong in
                     the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using  ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies)."
ORIGIN

  Query Match            47.8%;  Score 589.8;  DB 2;  Length 845;
  Best Local Similarity  97.8%;  Pred. No. 1.1e-117;
  Matches  609;  Conservative    0;  Mismatches   12;  Indels    2;  Gaps    1;

Qy          1 ATGCCAGGTGTCTGTGATAGGGCCCCTGACTTCCTCTCCCCGTCTGAAGACCAGGTGCTG 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        166 ATGCCAGGTGTCTGTGATAGGGCCCCTGACTTCCTCTCCCCGTCTGAAGACCAGGTGCTG 225

Qy         61 AGGCCTGCCTTGGGCAGCTCAGTGGCTCTGAACTGCACGGCTTGGGTAGTCTCTGGGCCC 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        226 AGGCCTGCCTTGGGCAGCTCAGTGGCTCTGAACTGCACGGCTTGGGTAGTCTCTGGGCCC 285

Qy        121 CACTGCTCCCTGCCTTCAGTCCAGTGGCTGAAAGACGGGCTTCCATTGGGAATTGGGGGC 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        286 CACTGCTCCCTGCCTTCAGTCCAGTGGCTGAAAGACGGGCTTCCATTGGGAATTGGGGGC 345

Qy        181 CACTACAGCCTCCACGAGTACTCCTGGGTCAAGGCCAACCTGTCAGAGGTGCTTGTGTCC 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        346 CACTACAGCCTCCACGAGTACTCCTGGGTCAAGGCCAACCTGTCAGAGGTGCTTGTGTCC 405

Qy        241 AGTGTCCTGGGGGTCAACGTGACCAGCACTGAAGTCTATGGGGCCTTCACCTGCTCCATC 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        406 AGTGTCCTGGGGGTCAACGTGACCAGCACTGAAGTCTATGGGGCCTTCACCTGCTCCATC 465

Qy        301 CAGAACATCAGCTTCTCCTCCTTCACTCTTCAGAGAGCTGGCCCTACAAGCCACGTGGCT 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        466 CAGAACATCAGCTTCTCCTCCTTCACTCTTCAGAGAGCTGGCCCTACAAGCCACGTGGCT 525

Qy        361 GCGGTGCTGGCCTCCCTCCTGGTCCTGCTGGCCCTGCTGCTGGCCGCCCTGCTCTATGTC 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        526 GCGGTGCTGGCCTCCCTCCTGGTCCTGCTGGCCCTGCTGCTGGCCGCCCTGCTCTATGTC 585

Qy        421 AAGTGCCGTCTCAACGTGCTGCTCTGGTACCAGGACGCGTATGGGGAGGTGGAGATAAAC 480
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        586 AAGTGCCGTCTCAACGTGCTGCTCTGGTACCAGGACGCGTATGGGGAGGTGGAGATAAAC 645

Qy        481 GACGGGAAGCTCTACGACGCCTACGTCTCCTACAGCGACTGCCCCGAGGACCGCAAGTTC 540
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        646 GACGGGAAGCTCTACGACGCCTACGTCTCCTACAGCGACTGCCCCGAGGACCGCAAGTTC 705

Qy        541 GTGAACTTCATCCTAAAGCCGCAGCTGGAGCGGCGT--CGGGGCTACAAGCTCTTCCTGG 598
              |||||||||||||||||||||||||||||||||||||   |||   ||||||||||||||
Db        706 GTGAACTTCATCCTAAAGCCGCAGCTGGAGCGGCGTTCGGGGCTAACAAGCTCTTCCTGG 765

Qy        599 ACGACCGCGACCTCCTGCCGCGC 621
              |||| || |||  |   | ||||
Db        766 ACGAACGGGACTCCTGCCGGCGC 788


RESULT 14
BX341382
LOCUS       BX341382                1003 bp    mRNA    linear   EST 07-APR-2004
DEFINITION  BX341382 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
            Homo sapiens cDNA clone CS0DJ015YD03 5-PRIME, mRNA sequence.
ACCESSION   BX341382
VERSION     BX341382.2  GI:46274994
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1003)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     On May 2, 2003 this sequence version replaced gi:30341987.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
FEATURES             Location/Qualifiers
     source          1. .1003
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DJ015YD03"
                     /cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
                     /cell_line="JURKAT"
                     /clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT
                     10-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN

  Query Match            47.5%;  Score 585.4;  DB 5;  Length 1003;
```

```
  Best Local Similarity   96.5%;  Pred. No. 9.7e-117;
  Matches  628;  Conservative    3;  Mismatches   14;  Indels    6;  Gaps    3;

Qy          1 ATGCCAGGTGTCTGTGATAGGGCCCCTGACTTCCTCTCCCCGTCTGAAGACCAGGTGCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        300 ATGCCAGGTGTCTGTGATAGGGCCCCTGACTTCCTCTCCCCGTCTGAAGACCAGGTGCTG 359

Qy         61 AGGCCTGCCTTGGGCAGCTCAGTGGCTCTGAACTGCACGGCTTGGGTAGTCTCTGGGCCC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        360 AGGCCTGCCTTGGGCAGCTCAGTGGCTCTGAACTGCACGGCTTGGGTAGTCTCTGGGCCC 419

Qy        121 CACTGCTCCCTGCCTTCAGTCCAGTGGCTGAAAGACGGGCTTCCATTGGGAATTGGGGGC 180
              ||||||||||||||||||||||||||||||||||:||||||||||||||||||||||||
Db        420 CACTGCTCCCTGCCTTCAGTCCAGTGGCTGAAAGAMGGGCTTCCATTGGGAATTGGGGGC 479

Qy        181 CACTACAGCCTCCACGAGTACTCCTGGGTCAAGGCCAACCTGTCAGAGGTGCTTGTGTCC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        480 CACTACAGCCTCCACGAGTACTCCTGGGTCAAGGCCAACCTGTCAGAGGTGCTTGTGTCC 539

Qy        241 AGTGTCCTGGGGGTCAACGTGACCAGCACTGAAGTCTATGGGGCCTTCACCTGCTCCATC 300
              |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db        540 AGTGTCCTGGGGGTCAACGTGACCAGCACTGAAGTCTATGGGG-CTTCACCTGCTCCATC 598

Qy        301 CAGAACATCAGCTTCTCCTCCTTCACTCTTCAGAGAGCTGGCCCTACAAGCCACGTGGCT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        599 CAGAACATCAGCTTCTCCTCCTTCACTCTTCAGAGAGCTGGCCCTACAAGCCACGTGGCT 658

Qy        361 GCGGTGCTGGCCTCCCTCCTGGTCCTGCTGGCCCTGCTGCTGGCCGCCCTGCTCTATGTC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        659 GCGGTGCTGGCCTCCCTCCTGGTCCTGCTGGCCCTGCTGCTGGCCGCCCTGCTCTATGTC 718

Qy        421 AAGTGCCGTCTCAACGTGCTGCTCTGGTACCAGGACGCGTATGGGGAGGTGGAGATAAAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        719 AAGTGCCGTCTCAACGTGCTGCTCTGGTACCAGGACGCGTATGGGGAGGTGGAGATAAAC 778

Qy        481 GACGGGAAGCTCTACGACGCCTACGTCTCCTACAGCGACTGCCCCGAGGACCGCAAGTTC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        779 GACGGGAAGCTCTACGACGCCTACGTCTCCTACAGCGACTGCCCCGAGGACCGCAAGTTC 838

Qy        541 GTGAACTTCATCCTAAAGCCGCAGCTGGAGCGGCGTCGGGGCTACAAGCTCTTCCTGGAC 600
              ||||||||||||||||||||||||||||||||||:||||||||| ||||||||||||||
Db        839 GTGAACTTCATCCTAAAGCCGCAGCTGGAGCGGCGTCSGGGCTACAA-CTCTTCCTGGAC 897

Qy        601 GACCGCGACCTCCTGCCGCGCGCTGAGCCCTCCGCCGACCTCTTGGTGAAC 651
              |||      | :|| | ||||||   |  |  |  ||||||||||||||||
Db        898 GAC----GCGAMCTCCTGCGCGCGCTGAGCCTCCGCGACCTCTTGGTGAAC 944


RESULT 15
BM714596
LOCUS       BM714596                584 bp    mRNA    linear   EST 28-FEB-2002
DEFINITION  UI-E-EJ0-ahr-n-24-0-UI.r1 UI-E-EJ0 Homo sapiens cDNA clone
            UI-E-EJ0-ahr-n-24-0-UI 5', mRNA sequence.
ACCESSION   BM714596
VERSION     BM714596.1  GI:19027854
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 584)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
  MEDLINE   97044477
   PUBMED   8889548
COMMENT     Contact: Soares, MB
            Coordinated Laboratory for Computational Genomics
            University of Iowa
            375 Newton Road , 4156  MEBRF, Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
            Email: bento-soares@uiowa.edu
            Tissue Procurement: Dr. Gregg Hageman
             cDNA Library preparation: Dr. M. Bento Soares, Univeristy of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, Univeristy of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, Univeristy of Iowa
             Clone Distribution: Researchers may obtain clones from Research
            Genetics (www.resgen.com).
             The following repetitive elements were found in this cDNA
            sequence: 35-104, >MER52#Unknown (matched compliment)
            Seq primer: M13 Reverse.
FEATURES              Location/Qualifiers
     source           1. .584
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /clone="UI-E-EJ0-ahr-n-24-0-UI"
                      /tissue_type="fetal eyes, lens, eye anterior segment,
                      optic nerve, retina, Retina Foveal and Macular, RPE and
                      Choroid"
                      /dev_stage="fetal and adult"
                      /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
                      /clone_lib="UI-E-EJ0"
                      /note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
                      modified polylinker; Site_1: EcoR I; Site_2: Not I;
                      UI-E-EJ0 is a subtracted cDNA library constructed
                      according to Bonaldo, Lennon and Soares, Genome Research,
                      6:791-806, 1996. First strand cDNA synthesis was primed
                      with an oligo-dT primer containing a Not I site. Double
                      stranded cDNA was ligated to an EcoR I adaptor, digested
                      with Not I, and cloned directionally into pT7T3-Pac
                      vector. The oligonucleotide used to prime the synthesis of
                      first-strand cDNA contains a library tag sequence that is
                      located between the Not I site and the (dT)18 tail. The
                      sequence tags for this library are: fetal eyes,
                      AGAATCAAGA; lens, CGATTAGCGA; eye anterior segment,
                      AATGCCGCAT; optic nerve, CCATTAAGTG; retina, CCGCG; Retina
                      Foveal and Macular, GTCC; RPE and Choroid, ACCTA. This
                      library was created for the program, Gene Discovery in the
                      Visual System, supported by National Eye Institute (NEI)."
ORIGIN

  Query Match             46.6%;  Score 575;  DB 4;  Length 584;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-114;
  Matches  575;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        659 GCTGCCGACGCCTCATCGTGGTGCTTTCGGACGCCTTCCTGAGCCGGGCCTGGTGCAGCC 718
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GCTGCCGACGCCTCATCGTGGTGCTTTCGGACGCCTTCCTGAGCCGGGCCTGGTGCAGCC 60

Qy        719 ACAGCTTCCGGGAGGGCCTGTGCCGGCTGCTGGAGCTCACCCGCAGACCCATCTTCATCA 778
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACAGCTTCCGGGAGGGCCTGTGCCGGCTGCTGGAGCTCACCCGCAGACCCATCTTCATCA 120

Qy        779 CCTTCGAGGGCCAGAGGCGCGACCCCGCGCACCCGGCGCTCCGCCTGCTGCGCCAGCACC 838
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CCTTCGAGGGCCAGAGGCGCGACCCCGCGCACCCGGCGCTCCGCCTGCTGCGCCAGCACC 180

Qy        839 GCCACCTGGTGACCTTGCTGCTCTGGAGGCCCGGCTCCGTGACTCCTTCCTCCGATTTTT 898
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCCACCTGGTGACCTTGCTGCTCTGGAGGCCCGGCTCCGTGACTCCTTCCTCCGATTTTT 240

Qy        899 GGAAAGAAGTGCAGCTGGCGCTGCCGCGGAAGGTGCGGTACAGGCCGGTGGAAGGAGACC 958
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GGAAAGAAGTGCAGCTGGCGCTGCCGCGGAAGGTGCGGTACAGGCCGGTGGAAGGAGACC 300

Qy        959 CCCAGACGCAGCTGCAGGACGACAAGGACCCCATGCTGATTCTTCGAGGCCGAGTCCCTG 1018
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CCCAGACGCAGCTGCAGGACGACAAGGACCCCATGCTGATTCTTCGAGGCCGAGTCCCTG 360

Qy       1019 AGGGCCGGGCCCTGGACTCAGAGGTGGACCCGGACCCTGAGGGCGACCTGGGTGTCCGGG 1078
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AGGGCCGGGCCCTGGACTCAGAGGTGGACCCGGACCCTGAGGGCGACCTGGGTGTCCGGG 420
```

```
Qy       1079 GGCCTGTTTTTGGAGAGCCATCAGCTCCACCGCACACCAGTGGGGTCTCGCTGGGAGAGA 1138
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGCCTGTTTTTGGAGAGCCATCAGCTCCACCGCACACCAGTGGGGTCTCGCTGGGAGAGA 480

Qy       1139 GCCGGAGCAGCGAAGTGGACGTCTCGGATCTCGGCTCGCGAAACTACAGTGCCCGCACAG 1198
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GCCGGAGCAGCGAAGTGGACGTCTCGGATCTCGGCTCGCGAAACTACAGTGCCCGCACAG 540

Qy       1199 ACTTCTACTGCCTGGTGTCCAAGGATGATATGTAG 1233
              |||||||||||||||||||||||||||||||||||
Db        541 ACTTCTACTGCCTGGTGTCCAAGGATGATATGTAG 575
```

Search completed: July 30, 2005, 01:21:04
Job time : 4602 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: August 1, 2005, 13:42:41 ; Search time 165 Seconds
(without alignments)
961.041 Million cell updates/sec

Title: US-09-598-443-2
Perfect score: 2147
Sequence: 1 MPGVCDRAPDFLSPSEDQVL.....GSRNYSARTDFYCLVSKDDM 410

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2147	100.0	410	2	AAY25426	Aay25426 Human SIG
2	2147	100.0	410	4	AAB61139	Aab61139 Human NOV
3	2147	100.0	410	6	ABO23241	Abo23241 Human bre
4	2147	100.0	410	7	ADM56381	Adm56381 Human cel
5	2147	100.0	410	8	ADF66745	Adf66745 Novel hum
6	2147	100.0	410	8	ADI19782	Adi19782 Human NOV
7	2147	100.0	410	8	ADO60255	Ado60255 Human NOV
8	2147	100.0	410	8	ADO20095	Ado20095 Human PRO
9	2147	100.0	410	8	ADO19087	Ado19087 Human PRO
10	2147	100.0	410	8	ADP54078	Adp54078 Human PRO
11	2147	100.0	410	8	ADP22970	Adp22970 PRO polyp
12	1876	87.4	504	5	AAU86151	Aau86151 Human PRO
13	1876	87.4	504	7	ADJ37329	Adj37329 Human tum
14	1876	87.4	504	8	ADG68253	Adg68253 Human PRO
15	1761	82.0	407	4	AAU17408	Aau17408 Novel sig
16	1761	82.0	407	4	AAM43652	Aam43652 Human pol
17	1761	82.0	407	7	ADB94116	Adb94116 Human nov
18	1761	82.0	407	8	ADM24673	Adm24673 Human PRO
19	1761	82.0	410	4	AAM43583	Aam43583 Human pol
20	1761	82.0	410	8	ADM24604	Adm24604 Human PRO
21	1749	81.5	363	3	AAB58416	Aab58416 Lung canc
22	1265	58.9	329	8	ADP46560	Adp46560 Human col
23	1265	58.9	361	8	ADP46559	Adp46559 Human col
24	1124	52.4	289	5	ADR41506	Adr41506 Human CD-
25	974.5	45.4	212	4	AAM25854	Aam25854 Human pro
26	631	29.4	164	8	ADP46561	Adp46561 Human col
27	542	25.2	113	5	ADK36203	Adk36203 Novel hum
28	267.5	12.5	658	2	AAY22164	Aay22164 Human TIG
29	267.5	12.5	686	2	AAY14128	Aay14128 Human IL-
30	267.5	12.5	686	5	AAE16567	Aae16567 Human int
31	267.5	12.5	686	5	ABP43986	Abp43986 Receptor
32	267.5	12.5	686	7	ADB31955	Adb31955 Human int
33	258.5	12.0	668	8	ABO84438	Abo84438 Mouse can
34	257.5	12.0	579	2	AAY14130	Aay14130 Human IL-
35	257.5	12.0	579	5	AAE16566	Aae16566 Human int
36	257.5	12.0	579	7	ADB31953	Adb31953 Human int
37	256.5	11.9	481	4	ABB10347	Abb10347 Human cDN
38	256.5	11.9	481	5	ABP66934	Abp66934 Human pol
39	256.5	11.9	486	2	AAY14129	Aay14129 Human IL-
40	256.5	11.9	486	5	AAE16575	Aae16575 Human int
41	256.5	11.9	486	7	ADB31971	Adb31971 Primate i
42	256.5	11.9	668	5	AAE16589	Aae16589 Human int
43	256.5	11.9	668	7	ADB31986	Adb31986 Primate i
44	256.5	11.9	696	3	AAY95298	Aay95298 Human Xre
45	256.5	11.9	696	4	AAM78808	Aam78808 Human pro

ALIGNMENTS

RESULT 1
AAY25426
ID AAY25426 standard; protein; 410 AA.
XX
AC AAY25426;
XX
DT 08-SEP-1999 (first entry)
XX
DE Human SIGIRR protein.
XX
KW SIGIRR; anti-inflammatory; anti-autoimmune disease; chromosome 11;
KW interleukin-1 receptor; 11p15.5; arthrogryposis multiplex congenita;
KW breast cancer; insulin-dependent diabetes; sickle cell anaemia;
KW bladder cancer; detection; cell-signal transduction; gene therapy;
KW immune regulation; cell proliferation; cell death; cell migration;
KW cell interaction; inflammation; marker; mass spectrometry; human.
XX
OS Homo sapiens.
XX
PN WO9932626-A1.
XX
PD 01-JUL-1999.
XX
PF 23-DEC-1998; 98WO-US027368.
XX
PR 23-DEC-1997; 97US-0068770P.
XX
PA (IMMV) IMMUNEX CORP.
XX
PI Sims JE;
XX
DR WPI; 1999-418925/35.
DR N-PSDB; AAX88091.
XX
PT New interleukin-1 receptor analog SIGIRR nucleic acid and proteins.
XX
PS Claim 2; Page 69-71; 72pp; English.
XX
CC This invention describes a novel human SIGIRR DNA, its allelic variants
CC or species homologues which have anti-inflammatory and anti-autoimmune
CC disease activity. SIGIRR is an analog of interleukin-1 receptor. The
CC products of the invention are used (a) as probes or primers for
CC identifying nucleic acid that encodes proteins with SIGIRR activity; (b)
CC to identify human chromosome 11, to map genes on this chromosome and to
CC identify disease-related genes (particularly in the region 11p15.5 where
CC genes are present associated with e.g. arthrogryposis multiplex
CC congenita, breast cancer, insulin-dependent diabetes, sickle cell
CC anaemia, bladder cancer), including detection of defective genes; (c) to

```
CC   study cell-signal transduction and the SIGIRR system, and (d) in gene
CC   therapy. Sense and antisense oligonucleotides can be used to inhibit
CC   expression of the SIGIRR gene. The proteins of the invention are used:
CC   (i) to study cellular processes (immune regulation, proliferation, death,
CC   migration, interaction with other cells and inflammation); (ii) to
CC   identify and purify proteins that associate with SIGIRR ligands and
CC   receptors, and to measure their biological activity; (iii) in screening
CC   for, and rational design of, potential inhibitors of activity; (iv)
CC   therapeutically against diseases mediated by SIGIRR polypeptide counter-
CC   structures; (v) as molecular weight (m.w.) markers in electrophoresis;
CC   (vi) for determining isoelectric points of unknown proteins; (vii) as
CC   controls for determining the extent of protein fragmentation (e.g. to aid
CC   characterization of protein structures by mass spectrometry); (viii) for
CC   generation of antibodies (Ab); (ix) to deliver diagnostic or therapeutic
CC   agents to cells that express SIGIRR binding molecules
XX
SQ   Sequence 410 AA;

  Query Match            100.0%;  Score 2147;  DB 2;  Length 410;
  Best Local Similarity  100.0%;  Pred. No. 1.2e-218;
  Matches  410;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MPGVCDRAPDFLSPSEDQVLRPALGSSVALNCTAWVVSGPHCSLPSVQWLKDGLPLGIGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MPGVCDRAPDFLSPSEDQVLRPALGSSVALNCTAWVVSGPHCSLPSVQWLKDGLPLGIGG 60

Qy         61 HYSLHEYSWVKANLSEVLVSSVLGVNVTSTEVYGAFTCSIQNISFSSFTLQRAGPTSHVA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 HYSLHEYSWVKANLSEVLVSSVLGVNVTSTEVYGAFTCSIQNISFSSFTLQRAGPTSHVA 120

Qy        121 AVLASLLVLLALLLAALLYVKCRLNVLLWYQDAYGEVEINDGKLYDAYVSYSDCPEDRKF 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AVLASLLVLLALLLAALLYVKCRLNVLLWYQDAYGEVEINDGKLYDAYVSYSDCPEDRKF 180

Qy        181 VNFILKPQLERRRGYKLFLDDRDLLPRAEPSADLLVNLSRCRRLIVVLSDAFLSRAWCSH 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VNFILKPQLERRRGYKLFLDDRDLLPRAEPSADLLVNLSRCRRLIVVLSDAFLSRAWCSH 240

Qy        241 SFREGLCRLLELTRRPIFITFEGQRRDPAHPALRLLRQHRHLVTLLLWRPGSVTPSSDFW 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SFREGLCRLLELTRRPIFITFEGQRRDPAHPALRLLRQHRHLVTLLLWRPGSVTPSSDFW 300

Qy        301 KEVQLALPRKVRYRPVEGDPQTQLQDDKDPMLILRGRVPEGRALDSEVDPDPEGDLGVRG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 KEVQLALPRKVRYRPVEGDPQTQLQDDKDPMLILRGRVPEGRALDSEVDPDPEGDLGVRG 360

Qy        361 PVFGEPSAPPHTSGVSLGESRSSEVDVSDLGSRNYSARTDFYCLVSKDDM 410
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 PVFGEPSAPPHTSGVSLGESRSSEVDVSDLGSRNYSARTDFYCLVSKDDM 410


RESULT 2
AAB61139
ID   AAB61139 standard; protein; 410 AA.
XX
AC   AAB61139;
XX
DT   30-MAR-2001  (first entry)
XX
DE   Human NOV9 protein.
XX
KW   Human; NOVX; antiinflammatory; cytostatic; neuroprotective;
KW   cerebroprotective; immunomodulator; vulnerary; vasotropic; gene therapy;
KW   hyperplasia; tumour; restenosis; psoriasis; Dupuytren's contracture;
KW   diabetes; rheumatoid arthritis; cerebral oedema; Alzheimer's disease.
XX
OS   Homo sapiens.
XX
PN   WO200075321-A2.
XX
PD   14-DEC-2000.
XX
PF   01-JUN-2000; 2000WO-US015303.
XX
PR   03-JUN-1999;   99US-0137322P.
PR   16-MAR-2000; 2000US-0189810P.
PR   22-MAR-2000; 2000US-0191158P.
PR   30-MAR-2000; 2000US-0193086P.
PR   31-MAY-2000; 2000US-00137322.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Fernandes E,  Herrman J,  Vernet C;
XX
DR   WPI; 2001-102403/11.
DR   N-PSDB; AAF27857.
XX
PT   New NOVX polypeptides and polynucleotides, useful in gene therapy, as a
PT   diagnostic marker, protein therapeutic, antibody or small molecule drug
PT   target for treating immune, proliferative and metabolic diseases and
PT   wound healing.
XX
PS   Claim 1; Page 36-38; 194pp; English.
XX
CC   The present sequence is a new isolated polypeptide (NOVX). The NOVX
CC   polypeptides, NOVX nucleic acids, and anti-NOVX antibodies are useful for
CC   treating or preventing NOVX-associated disorders. They are also useful
CC   for determining the presence of or a predisposition to a disease
CC   associated with altered levels of the NOVX polypeptide or nucleic acid.
CC   These NOVX-associated disorders include hyperplasias, tumours,
CC   restenosis, psoriasis, Dupuytren's contracture, diabetic complications,
CC   rheumatoid arthritis, cerebral lesions, diabetic neuropathies, cerebral
CC   oedema, senile dementia or Alzheimer's disease. The NOVX polynucleotides
CC   are especially useful in gene therapy. Specifically, NOVX is useful as a
CC   diagnostic marker or prognostic marker, protein therapeutic and antibody
CC   target or small molecule drug target to treat disorders in the immune
CC   response pathway, thyroid and metabolic diseases, bone metabolic
CC   disorders, diseases of the pancreas (e.g. diabetes or digestive
CC   disorders), proliferative diseases, or tissue regeneration and
CC   development (e.g. wound healing or treatment of burns)
XX
SQ   Sequence 410 AA;

  Query Match            100.0%;  Score 2147;  DB 4;  Length 410;
  Best Local Similarity  100.0%;  Pred. No. 1.2e-218;
  Matches  410;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MPGVCDRAPDFLSPSEDQVLRPALGSSVALNCTAWVVSGPHCSLPSVQWLKDGLPLGIGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MPGVCDRAPDFLSPSEDQVLRPALGSSVALNCTAWVVSGPHCSLPSVQWLKDGLPLGIGG 60

Qy         61 HYSLHEYSWVKANLSEVLVSSVLGVNVTSTEVYGAFTCSIQNISFSSFTLQRAGPTSHVA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 HYSLHEYSWVKANLSEVLVSSVLGVNVTSTEVYGAFTCSIQNISFSSFTLQRAGPTSHVA 120

Qy        121 AVLASLLVLLALLLAALLYVKCRLNVLLWYQDAYGEVEINDGKLYDAYVSYSDCPEDRKF 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AVLASLLVLLALLLAALLYVKCRLNVLLWYQDAYGEVEINDGKLYDAYVSYSDCPEDRKF 180

Qy        181 VNFILKPQLERRRGYKLFLDDRDLLPRAEPSADLLVNLSRCRRLIVVLSDAFLSRAWCSH 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VNFILKPQLERRRGYKLFLDDRDLLPRAEPSADLLVNLSRCRRLIVVLSDAFLSRAWCSH 240

Qy        241 SFREGLCRLLELTRRPIFITFEGQRRDPAHPALRLLRQHRHLVTLLLWRPGSVTPSSDFW 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SFREGLCRLLELTRRPIFITFEGQRRDPAHPALRLLRQHRHLVTLLLWRPGSVTPSSDFW 300

Qy        301 KEVQLALPRKVRYRPVEGDPQTQLQDDKDPMLILRGRVPEGRALDSEVDPDPEGDLGVRG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 KEVQLALPRKVRYRPVEGDPQTQLQDDKDPMLILRGRVPEGRALDSEVDPDPEGDLGVRG 360

Qy        361 PVFGEPSAPPHTSGVSLGESRSSEVDVSDLGSRNYSARTDFYCLVSKDDM 410
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 PVFGEPSAPPHTSGVSLGESRSSEVDVSDLGSRNYSARTDFYCLVSKDDM 410
```

RESULT 3
ABO23241
ID ABO23241 standard; protein; 410 AA.
XX
AC ABO23241;
XX
DT 03-SEP-2003 (first entry)
XX
DE Human breast tumour associated protein 47-like polypeptide NOV9.
XX
KW Tissue typing; cancer; breast cancer; colon cancer; lung cancer; sarcoma;
KW pancreatic cancer; uterine cancer; organ transplantation disorder;
KW cardiovascular disease; melanoma; atherosclerosis; diabetes mellitus;
KW ischaemic heart disease; haemorrhage; peripheral vascular disease;
KW thrombosis; hypertension; systemic lupus erythematosus; haematopoiesis;
KW tissue regeneration; wound healing; hyperproliferative disorder;
KW psoriasis; neural disorder; Parkinson's disease; Alzheimer's disease;
KW Huntington's disease; multiple sclerosis; amyotrophic lateral sclerosis;
KW ALS; peripheral neuropathy; nervous system tumour; neurotoxin; tremor;
KW neuropathy; acute brain injury; peripheral nerve trauma; human; NOVX;
KW gene therapy; epilepsy; breast tumour associated protein 47.
XX
OS Homo sapiens.
XX
PN US2003027158-A1.
XX
PD 06-FEB-2003.
XX
PF 15-OCT-2001; 2001US-00977418.
XX
PR 03-JUN-1999; 99US-0137322P.
PR 16-MAR-2000; 2000US-0189810P.
PR 22-MAR-2000; 2000US-0191158P.
PR 30-MAR-2000; 2000US-0193086P.
PR 03-MAY-2000; 2000US-0201388P.
PR 31-MAY-2000; 2000US-00584411.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Fernandes E, Herrman J, Vernet C;
XX
DR WPI; 2003-492028/46.
DR N-PSDB; ACD40262.
XX
PT New nucleic acid sequence encoding a human breast tumor-associated
PT protein 47-like polypeptide, useful for treating cardiovascular
PT disorders, neural disorders, diabetes mellitus and cancers.
XX
PS Disclosure; Page 24-25; 100pp; English.
XX
CC The invention relates to a new isolated NOV4 nucleic acid. The nucleic
CC acid is useful for identifying a compound that binds the nucleic acid.
CC The nucleic acid is useful in gene therapy, in screening assays, in
CC detection assays e.g. chromosomal mapping, cell and tissue typing and
CC forensic biology, predictive medicine e.g. diagnostic assays, prognostic
CC assays, monitoring clinical trials and pharmacogenomics and methods of
CC treatment including therapeutic and prophylactic. The nucleic acid is
CC also useful for expressing NOVX protein. The nucleic acid is also useful
CC to provide polynucleotide reagents e.g. labelled probes that are useful
CC in an in situ hybridisation technique, for identifying a specific tissue
CC (for example brain tissue) and for use in forensic science. The nucleic
CC acid is also useful for mapping genes on a chromosome and thus locating
CC gene regions associated with genetic disease, identifying an individual
CC from a minute biological sample and to aid in forensic identification of
CC biological sample. The nucleic acid is also useful for treating cancer,
CC especially cancers of the breast, colon, lung, pancreas or uterus, or a
CC melanoma or sarcoma. The nucleic acid is also useful for treating
CC disorders related to organ transplantation, cardiovascular diseases,
CC atherosclerosis, ischaemic heart disease, haemorrhage, diabetes mellitus,
CC peripheral vascular disease, thrombosis, hypertension and systemic lupus
CC erythematosus. NOVX protein encoded by the nucleic acid is useful for
CC regulating haematopoiesis, for regeneration of bone, cartilage, tendon
CC ligament and/or nerve tissue growth or regeneration and for wound
CC healing. The nucleic acid is also useful for treating infections,
CC hyperproliferative disorders e.g. psoriasis, and neural disorders
CC including Parkinson's disease, Alzheimer's disease, Huntington's disease,
CC multiple sclerosis, amyotrophic lateral sclerosis (ALS), peripheral
CC neuropathy, tumours of the nervous system, exposure to neurotoxins, acute
CC brain injury, peripheral nerve trauma or injury and other neuropathies,
CC epilepsy, and/or tremors. The present sequence represents the amino acid
CC sequence of a human breast tumour associated protein 47-like polypeptide
XX
SQ Sequence 410 AA;

Query Match 100.0%; Score 2147; DB 6; Length 410;
Best Local Similarity 100.0%; Pred. No. 1.2e-218;
Matches 410; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MPGVCDRAPDFLSPSEDQVLRPALGSSVALNCTAWVVSGPHCSLPSVQWLKDGLPLGIGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MPGVCDRAPDFLSPSEDQVLRPALGSSVALNCTAWVVSGPHCSLPSVQWLKDGLPLGIGG 60

Qy         61 HYSLHEYSWVKANLSEVLVSSVLGVNVTSTEVYGAFTCSIQNISFSSFTLQRAGPTSHVA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 HYSLHEYSWVKANLSEVLVSSVLGVNVTSTEVYGAFTCSIQNISFSSFTLQRAGPTSHVA 120

Qy        121 AVLASLLVLLALLLAALLYVKCRLNVLLWYQDAYGEVEINDGKLYDAYVSYSDCPEDRKF 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AVLASLLVLLALLLAALLYVKCRLNVLLWYQDAYGEVEINDGKLYDAYVSYSDCPEDRKF 180

Qy        181 VNFILKPQLERRRGYKLFLDDRDLLPRAEPSADLLVNLSRCRRLIVVLSDAFLSRAWCSH 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VNFILKPQLERRRGYKLFLDDRDLLPRAEPSADLLVNLSRCRRLIVVLSDAFLSRAWCSH 240

Qy        241 SFREGLCRLLELTRRPIFITFEGQRRDPAHPALRLLRQHRHLVTLLLWRPGSVTPSSDFW 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SFREGLCRLLELTRRPIFITFEGQRRDPAHPALRLLRQHRHLVTLLLWRPGSVTPSSDFW 300

Qy        301 KEVQLALPRKVRYRPVEGDPQTQLQDDKDPMLILRGRVPEGRALDSEVDPDPEGDLGVRG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 KEVQLALPRKVRYRPVEGDPQTQLQDDKDPMLILRGRVPEGRALDSEVDPDPEGDLGVRG 360

Qy        361 PVFGEPSAPPHTSGVSLGESRSSEVDVSDLGSRNYSARTDFYCLVSKDDM 410
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 PVFGEPSAPPHTSGVSLGESRSSEVDVSDLGSRNYSARTDFYCLVSKDDM 410
```

RESULT 4
ADM56381
ID ADM56381 standard; protein; 410 AA.
XX
AC ADM56381;
XX
DT 03-JUN-2004 (first entry)
XX
DE Human cell adhesion molecule NOV9.
XX
KW Human; cell adhesion molecule; NOVX; cancer; leukaemia; lymphoma;
KW melanoma; neurological disorder; epilepsy;
KW ischaemic cerebrovascular disease; stroke; Alzheimer's disease;
KW Pick's disease; vesicular transport disorder; cystic fibrosis;
KW diabetes mellitus; Grave's disease; goiter; gastrointestinal disorder;
KW ulcerative colitis; gastric ulcer; duodenal disorder; autoimmune disease;
KW allergic reaction; autoimmune haemolytic anaemia; rheumatoid arthritis;
KW viral infection; bacterial infection; fungal infection;
KW helminthic infection; protozoal infections.
XX
OS Homo sapiens.
XX
PN US2003082554-A1.
XX
PD 01-MAY-2003.
XX
PF 15-OCT-2001; 2001US-00977033.
XX
PR 03-JUN-1999; 99US-0137322P.

PR 16-MAR-2000; 2000US-0189810P.
PR 22-MAR-2000; 2000US-0191158P.
PR 30-MAR-2000; 2000US-0193086P.
PR 03-MAY-2000; 2000US-0201388P.
PR 31-MAY-2000; 2000US-00584411.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Fernandes E, Herrman J, Vernet C;
XX
DR WPI; 2003-616079/58.
DR N-PSDB; ADM56382.
XX
PT New nucleic acids encoding human cell adhesion molecule-like proteins,
PT useful for treating e.g. cancers, neurological disorders, viral,
PT bacterial, fungal, helminthic and protozoal infections.
XX
PS Disclosure; SEQ ID NO 18; 78pp; English.
XX
CC The invention relates to an isolated nucleic acid encoding a human cell
CC adhesion molecule-like protein, comprising a sequence encoding a
CC polypeptide having a sequence appearing as ADM56387, a sequence at least
CC 90% identical to the nucleic acid, a sequence encoding a polypeptide
CC having conservative amino acid substitutions to the protein or a fragment
CC comprising at least 20 nucleotides. Also included are an oligonucleotide
CC sequence that is complementary to (and hybridises under stringent
CC conditions with) the nucleic acid (or a portion of it), a vector
CC comprising the nucleic acid, a cell comprising the vector, a
CC pharmaceutical composition comprising the nucleic acid and a
CC pharmaceutical carrier, a process for producing the polypeptide, a
CC process for identifying a compound that binds the nucleic acid, and a
CC compound identified by the process. Disclosed as new are the cDNA and
CC proteins for novel cell adhesion molecules (termed NOVX, being NOV1-23).
CC The NOVX polypeptide, nucleic acid or antibody are useful in the
CC manufacture of a medicament for treating a syndrome associated with a
CC human disease selected from NOVX-associated disorder, such as cancers
CC (e.g. leukaemia, lymphoma, melanoma or cancer of the liver, lung, muscle,
CC ovary, testis and uterus), neurological disorders (e.g. epilepsy,
CC ischaemic cerebrovascular disease, stroke, Alzheimer's disease or Pick's
CC disease), disorders of vesicular transport (e.g. cystic fibrosis,
CC diabetes mellitus, Grave's disease, or goiter), gastrointestinal
CC disorders (e.g. ulcerative colitis, or gastric and duodenal disorders),
CC autoimmune diseases (e.g. allergic reactions, autoimmune haemolytic
CC anaemia, or rheumatoid arthritis), viral, bacterial, fungal, helminthic
CC and protozoal infections. The polypeptides can be used as immunogens to
CC produce antibodies and as vaccines. The sequences may further be used in
CC chromosome mapping, identifying individual from minute biological samples
CC (tissue typing), and in forensic identification of a biological sample.
CC NOTE: The authors have mis-labelled the sequences as they appear on pages
CC 12-28 of the patent, it is clear from table 3, the examples and the
CC claims that the SEQ ID numbers for the cDNAs should be the odd numbers
CC from 1-45 and the proteins should be the even numbers from 2-46. The
CC present sequence represents a NOVX cell adhesion molecule of the
CC invention.
XX
SQ Sequence 410 AA;

Query Match 100.0%; Score 2147; DB 7; Length 410;
Best Local Similarity 100.0%; Pred. No. 1.2e-218;
Matches 410; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MPGVCDRAPDFLSPSEDQVLRPALGSSVALNCTAWVVSGPHCSLPSVQWLKDGLPLGIGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MPGVCDRAPDFLSPSEDQVLRPALGSSVALNCTAWVVSGPHCSLPSVQWLKDGLPLGIGG 60

Qy         61 HYSLHEYSWVKANLSEVLVSSVLGVNVTSTEVYGAFTCSIQNISFSSFTLQRAGPTSHVA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 HYSLHEYSWVKANLSEVLVSSVLGVNVTSTEVYGAFTCSIQNISFSSFTLQRAGPTSHVA 120

Qy        121 AVLASLLVLLALLLAALLYVKCRLNVLLWYQDAYGEVEINDGKLYDAYVSYSDCPEDRKF 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AVLASLLVLLALLLAALLYVKCRLNVLLWYQDAYGEVEINDGKLYDAYVSYSDCPEDRKF 180

Qy        181 VNFILKPQLERRRGYKLFLDDRDLLPRAEPSADLLVNLSRCRRLIVVLSDAFLSRAWCSH 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VNFILKPQLERRRGYKLFLDDRDLLPRAEPSADLLVNLSRCRRLIVVLSDAFLSRAWCSH 240

Qy        241 SFREGLCRLLELTRRPIFITFEGQRRDPAHPALRLLRQHRHLVTLLLWRPGSVTPSSDFW 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SFREGLCRLLELTRRPIFITFEGQRRDPAHPALRLLRQHRHLVTLLLWRPGSVTPSSDFW 300

Qy        301 KEVQLALPRKVRYRPVEGDPQTQLQDDKDPMLILRGRVPEGRALDSEVDPDPEGDLGVRG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 KEVQLALPRKVRYRPVEGDPQTQLQDDKDPMLILRGRVPEGRALDSEVDPDPEGDLGVRG 360

Qy        361 PVFGEPSAPPHTSGVSLGESRSSEVDVSDLGSRNYSARTDFYCLVSKDDM 410
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 PVFGEPSAPPHTSGVSLGESRSSEVDVSDLGSRNYSARTDFYCLVSKDDM 410
```

RESULT 5
ADF66745
ID ADF66745 standard; protein; 410 AA.
XX
AC ADF66745;
XX
DT 12-FEB-2004 (first entry)
XX
DE Novel human protein NOV9.
XX
KW cytostatic; hepatotropic; vulnerary; antipsoriatic; osteopathic;
KW antiarthritic; antiatherosclerotic; haemostatic; vasotropic;
KW thrombolytic; antidiabetic; hypotensive; dermatalogical;
KW immunosuppressive; antiinflammatory; immunostimulant; fungicide;
KW virucide; protozoacide; neuroprotective; antirheumatic; antiarthritic;
KW antiasthmatic; antiparkinsonian; nootropic; anticonvulsant;
KW NOVX modulator; cancer; hyperproliferative disease; cirrhosis; keloid;
KW psoriasis; tissue hypertrophy; osteoarthritis;
KW atherosclerotic plaque formation; haemorrhage; ischaemic disease;
KW thrombosis; diabetes mellitus; hypertension; hypothyroidism;
KW immune deficiency; severe combined immunodeficiency; SCID; infection;
KW malaria; candidiasis; autoimmune disorder; connective tissue disease;
KW multiple sclerosis; systemic lupus, erythematosus; rheumatoid arthritis;
KW autoimmune pulmonary inflammation; Guillain-Barre syndrome;
KW autoimmune thyroiditis; insulin dependent diabetes mellitus;
KW myasthenia gravis; graft- versus-host disease;
KW autoimmune inflammatory eye disease; asthma; haematopoiesis;
KW tissue regeneration; wound healing; tissue repair; burn; incision; ulcer;
KW periodontal disease; Alzhemier's disease; Parkinson's disease;
KW Huntington's disease; amyotrophic lateral sclerosis; Shy-Drager syndrome;
KW human.
XX
OS Homo sapiens.
XX
PN US2003199103-A1.
XX
PD 23-OCT-2003.
XX
PF 15-OCT-2001; 2001US-00977639.
XX
PR 03-JUN-1999; 99US-0137322P.
PR 16-MAR-2000; 2000US-0189810P.
PR 22-MAR-2000; 2000US-0191158P.
PR 30-MAR-2000; 2000US-0193086P.
PR 03-MAY-2000; 2000US-0201388P.
PR 31-MAY-2000; 2000US-00584411.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Fernandes E, Herrman J, Vernet C;
XX
DR WPI; 2004-021196/02.
DR N-PSDB; ADF66744.
XX
PT Novel substantially NOVX polypeptide useful for diagnosing, preventing
PT and treating diseases e.g., cancer, multiple sclerosis, systemic lupus

```
PT   erythematous.
XX
PS   Disclosure; SEQ ID NO 18; 165pp; English.
XX
CC   The invention describes a substantially purified polypeptide (I) having
CC   amino acid sequence chosen from a fully defined NOVX sequence (S1) of 708
CC   amino acids as given in the specification, or polypeptide having one or
CC   more conservative amino acid substitutions of (S1), or mutant or variant
CC   of (S1). (I) having (S1) is useful for diagnosing a pathological
CC   condition associated with (I) or its activity in a subject e.g. cancer.
CC   (I) useful in treatment of cancer, hyperproliferative diseases,
CC   cirrhosis, keloid, psoriasis, tissue hypertrophy, osteoarthritis,
CC   atherosclerotic plaque formation, haemorrhage, ischaemic heart or renal
CC   disease, thrombosis, diabetes mellitus, hypertension, hypothyroidism. (I)
CC   is useful in treatment of various immune deficiencies and disorders such
CC   as severe combined immunodeficiency (SCID), bacterial infection, viral
CC   infection such as herpes viral infection, protozoan infection such as
CC   malaria, fungal infection such as candidiasis. (I) is also useful in
CC   treating autoimmune disorders such as connective tissue disease, multiple
CC   sclerosis, systemic lupus, erythematosus, rheumatoid arthritis,
CC   autoimmune pulmonary inflammation, Guillain-Barre syndrome, autoimmune
CC   thyroiditis, insulin dependent diabetes mellitus, myasthenia gravis,
CC   graft- versus-host disease and autoimmune inflammatory eye disease and
CC   asthma. (I) useful in regulation of haematopoiesis, regeneration and
CC   tissue growth of bone, cartilage, tendon, ligament and useful for wound
CC   healing and tissue repair. (I) is also useful in treatment of burns,
CC   incisions and ulcers. (I) also useful in treatment of periodontal
CC   disease, Alzhemier's disease, Parkinson's disease, Huntington's disease,
CC   amyotrophic lateral sclerosis, and Shy-Drager syndrome. (I) has effective
CC   antitumour and antiinflammatory activity. This is the amino acid sequence
CC   of a novel human NOVX protein.
XX
SQ   Sequence 410 AA;

  Query Match             100.0%;  Score 2147;  DB 8;  Length 410;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-218;
  Matches  410;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MPGVCDRAPDFLSPSEDQVLRPALGSSVALNCTAWVVSGPHCSLPSVQWLKDGLPLGIGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MPGVCDRAPDFLSPSEDQVLRPALGSSVALNCTAWVVSGPHCSLPSVQWLKDGLPLGIGG 60

Qy         61 HYSLHEYSWVKANLSEVLVSSVLGVNVTSTEVYGAFTCSIQNISFSSFTLQRAGPTSHVA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 HYSLHEYSWVKANLSEVLVSSVLGVNVTSTEVYGAFTCSIQNISFSSFTLQRAGPTSHVA 120

Qy        121 AVLASLLVLLALLLAALLYVKCRLNVLLWYQDAYGEVEINDGKLYDAYVSYSDCPEDRKF 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AVLASLLVLLALLLAALLYVKCRLNVLLWYQDAYGEVEINDGKLYDAYVSYSDCPEDRKF 180

Qy        181 VNFILKPQLERRRGYKLFLDDRDLLPRAEPSADLLVNLSRCRRLIVVLSDAFLSRAWCSH 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VNFILKPQLERRRGYKLFLDDRDLLPRAEPSADLLVNLSRCRRLIVVLSDAFLSRAWCSH 240

Qy        241 SFREGLCRLLELTRRPIFITFEGQRRDPAHPALRLLRQHRHLVTLLLWRPGSVTPSSDFW 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SFREGLCRLLELTRRPIFITFEGQRRDPAHPALRLLRQHRHLVTLLLWRPGSVTPSSDFW 300

Qy        301 KEVQLALPRKVRYRPVEGDPQTQLQDDKDPMLILRGRVPEGRALDSEVDPDPEGDLGVRG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 KEVQLALPRKVRYRPVEGDPQTQLQDDKDPMLILRGRVPEGRALDSEVDPDPEGDLGVRG 360

Qy        361 PVFGEPSAPPHTSGVSLGESRSSEVDVSDLGSRNYSARTDFYCLVSKDDM 410
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 PVFGEPSAPPHTSGVSLGESRSSEVDVSDLGSRNYSARTDFYCLVSKDDM 410


RESULT 6
ADI19782
ID   ADI19782 standard; protein; 410 AA.
XX
AC   ADI19782;
XX
DT   22-APR-2004  (first entry)
XX
DE   Human NOV9 protein.
XX
KW   Secreted protein; NOVX; diagnosis; metabolic disorder; diabetes; obesity;
KW   infection; anorexia; cancer; cardiovascular disease; hypertension;
KW   atherosclerosis; neurodegenerative disorder; Alzheimer's disease;
KW   Parkinson's disease; epilepsy; immune disorder; osteoarthritis;
KW   haematopoietic disorder; inflammatory skin disorder; asthma;
KW   dyslipidemia; neurogenesis; cell differentiation; cell proliferation;
KW   haematopoiesis; wound healing; angiogenesis; chromosome mapping;
KW   tissue typing; preventive medicine; pharmacogenomic; gene therapy;
KW   anorectic; cardiant; virucide; antibacterial; fungicide; protozoacide;
KW   nootropic; neuroprotective; dermatological; human.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 37
FT                   /note= "Encoded by NTC"
FT   Domain          138. .410
FT                   /note = Extracellular domain
XX
PN   US2004002134-A1.
XX
PD   01-JAN-2004.
XX
PF   15-OCT-2001; 2001US-00977819.
XX
PR   03-JUN-1999;   99US-0137322P.
PR   16-MAR-2000; 2000US-0189810P.
PR   22-MAR-2000; 2000US-0191158P.
PR   30-MAR-2000; 2000US-0193086P.
PR   03-MAY-2000; 2000US-0201388P.
PR   31-MAY-2000; 2000US-00584411.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Fernandes ER,  Herrman JL,  Vernet CAM;
XX
DR   WPI; 2004-070737/07.
DR   N-PSDB; ADI19781.
XX
PT   New NOVX nucleic acids encoding human KIAA0768 protein-like and human
PT   protein PRO-228 polypeptides, useful for treating NOVX-associated
PT   disorders.
XX
PS   Disclosure; SEQ ID NO 18; 95pp; English.
XX
CC   The present invention is based in part on the discovery of novel secreted
CC   and membrane-bound polypeptides and their encoding polynucleotides. The
CC   nucleic acids and polypeptides are collectively referred as NOVX. The
CC   invention is useful for treating, preventing and diagnosing diseases such
CC   as metabolic disorders, diabetes, obesity, infectious diseases such as
CC   viral, bacterial, fungal, helminthic and protozoal infections, anorexia,
CC   cancer, cardiovascular diseases such as hypertension and atherosclerosis,
CC   neurodegenerative disorders, Alzheimer's disease, Parkinson's disease,
CC   epilepsy, immune disorders such as osteoarthritis, haematopoietic
CC   disorders, inflammatory skin disorders, asthma and various dyslipidemias.
CC   The invention is also useful as targets for the identification of small
CC   molecules that modulate or inhibit e.g. neurogenesis, cell
CC   differentiation, cell proliferation, haematopoiesis, wound healing and
CC   angiogenesis, as hybridisation probes, in chromosome mapping, tissue
CC   typing, preventive medicine and pharmacogenomics. The invention is also
CC   useful in gene therapy. The present sequence is human NOV protein.
XX
SQ   Sequence 410 AA;

  Query Match             100.0%;  Score 2147;  DB 8;  Length 410;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-218;
  Matches  410;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MPGVCDRAPDFLSPSEDQVLRPALGSSVALNCTAWVVSGPHCSLPSVQWLKDGLPLGIGG 60
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MPGVCDRAPDFLSPSEDQVLRPALGSSVALNCTAWVVSGPHCSLPSVQWLKDGLPLGIGG 60

Qy         61 HYSLHEYSWVKANLSEVLVSSVLGVNVTSTEVYGAFTCSIQNISFSSFTLQRAGPTSHVA 120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 HYSLHEYSWVKANLSEVLVSSVLGVNVTSTEVYGAFTCSIQNISFSSFTLQRAGPTSHVA 120

Qy        121 AVLASLLVLLALLLAALLYVKCRLNVLLWYQDAYGEVEINDGKLYDAYVSYSDCPEDRKF 180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AVLASLLVLLALLLAALLYVKCRLNVLLWYQDAYGEVEINDGKLYDAYVSYSDCPEDRKF 180

Qy        181 VNFILKPQLERRRGYKLFLDDRDLLPRAEPSADLLVNLSRCRRLIVVLSDAFLSRAWCSH 240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VNFILKPQLERRRGYKLFLDDRDLLPRAEPSADLLVNLSRCRRLIVVLSDAFLSRAWCSH 240

Qy        241 SFREGLCRLLELTRRPIFITFEGQRRDPAHPALRLLRQHRHLVTLLLWRPGSVTPSSDFW 300
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SFREGLCRLLELTRRPIFITFEGQRRDPAHPALRLLRQHRHLVTLLLWRPGSVTPSSDFW 300

Qy        301 KEVQLALPRKVRYRPVEGDPQTQLQDDKDPMLILRGRVPEGRALDSEVDPDPEGDLGVRG 360
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 KEVQLALPRKVRYRPVEGDPQTQLQDDKDPMLILRGRVPEGRALDSEVDPDPEGDLGVRG 360

Qy        361 PVFGEPSAPPHTSGVSLGESRSSEVDVSDLGSRNYSARTDFYCLVSKDDM 410
                |||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 PVFGEPSAPPHTSGVSLGESRSSEVDVSDLGSRNYSARTDFYCLVSKDDM 410
```

RESULT 7
ADO60255
ID AD060255 standard; protein; 410 AA.
XX
AC AD060255;
XX
DT 15-JUL-2004 (first entry)
XX
DE Human NOV9 protein.
XX
KW Human; NOVX protein; cancer; hyperproliferative disease; cirrhosis;
KW keloid; psoriasis; tissue hypertrophy; osteoarthritis;
KW atherosclerotic plaque formation; haemorrhage; ischaemic heart disease;
KW renal disease; thrombosis; diabetes mellitus; hypertension;
KW hypothyroidism; severe combined immunodeficiency; SCID; infection;
KW malaria; candidiasis; autoimmune disorder; connective tissue disease;
KW multiple sclerosis; systemic lupus erythematosus; rheumatoid arthritis;
KW autoimmune pulmonary inflammation; Guillain-Barre syndrome;
KW autoimmune thyroiditis; myasthenia gravis; graft- versus-host disease;
KW autoimmune inflammatory eye disease; asthma; burn; incision; ulcer;
KW periodontal disease; Alzhemier's disease; Parkinson's disease;
KW Huntington's disease; amyotrophic lateral sclerosis; Shy-Drager syndrome;
KW haematopoiesis; wound healing; tissue repair; antitumor;
KW antiinflammatory.
XX
OS Homo sapiens.
XX
PN US2003134430-A1.
XX
PD 17-JUL-2003.
XX
PF 15-OCT-2001; 2001US-00977751.
XX
PR 03-JUN-1999; 99US-0137322P.
PR 16-MAR-2000; 2000US-0189810P.
PR 22-MAR-2000; 2000US-0191158P.
PR 30-MAR-2000; 2000US-0193086P.
PR 03-MAY-2000; 2000US-0201388P.
PR 31-MAY-2000; 2000US-00584411.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Fernandes E, Herrman J, Vernet C;
XX
DR WPI; 2004-068928/07.
DR N-PSDB; AD060254.
XX
PT Novel substantially purified NOVX polypeptide for treating severe
PT combined immunodeficiency, candidiasis, cancer, asthma, multiple
PT sclerosis, systemic lupus erythematosus.
XX
PS Disclosure; SEQ ID NO 18; 155pp; English.
XX
CC The invention relates to human NOVX polypeptides and polynucleotides.
CC NOVX sequences are useful in the treatment of cancer, hyperproliferative
CC diseases, cirrhosis, keloid, psoriasis, tissue hypertrophy,
CC osteoarthritis, atherosclerotic plaque formation, haemorrhage, ischaemic
CC heart or renal disease, thrombosis, diabetes mellitus, hypertension,
CC hypothyroidism, asthma, burns, incisions, ulcers, periodontal disease,
CC Alzhemier's disease, Parkinson's disease, Huntington's disease,
CC amyotrophic lateral sclerosis, Shy-Drager syndrome, immune deficiencies
CC and disorders such as severe combined immunodeficiency (SCID), bacterial
CC infection, viral infection e.g. herpes viral infection, protozoan
CC infection e.g. malaria, fungal infection e.g. candidiasis, autoimmune
CC disorders such as connective tissue disease, multiple sclerosis, systemic
CC lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary
CC inflammation, Guillain-Barre syndrome, autoimmune thyroiditis, insulin
CC dependent diabetes mellitus, myasthenia gravis, graft-versus-host disease
CC and autoimmune inflammatory eye disease. The invention is useful in
CC regeneration and tissue growth of bone, cartilage, tendon, ligament,
CC haematopoiesis regulation, wound healing and tissue repair. Sequences of
CC the invention also exhibit antitumor and antiinflammatory activities. The
CC present sequence is human NOVX protein.
XX
SQ Sequence 410 AA;

```
  Query Match             100.0%;  Score 2147;  DB 8;  Length 410;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-218;
  Matches  410;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MPGVCDRAPDFLSPSEDQVLRPALGSSVALNCTAWVVSGPHCSLPSVQWLKDGLPLGIGG 60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MPGVCDRAPDFLSPSEDQVLRPALGSSVALNCTAWVVSGPHCSLPSVQWLKDGLPLGIGG 60

Qy         61 HYSLHEYSWVKANLSEVLVSSVLGVNVTSTEVYGAFTCSIQNISFSSFTLQRAGPTSHVA 120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 HYSLHEYSWVKANLSEVLVSSVLGVNVTSTEVYGAFTCSIQNISFSSFTLQRAGPTSHVA 120

Qy        121 AVLASLLVLLALLLAALLYVKCRLNVLLWYQDAYGEVEINDGKLYDAYVSYSDCPEDRKF 180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AVLASLLVLLALLLAALLYVKCRLNVLLWYQDAYGEVEINDGKLYDAYVSYSDCPEDRKF 180

Qy        181 VNFILKPQLERRRGYKLFLDDRDLLPRAEPSADLLVNLSRCRRLIVVLSDAFLSRAWCSH 240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VNFILKPQLERRRGYKLFLDDRDLLPRAEPSADLLVNLSRCRRLIVVLSDAFLSRAWCSH 240

Qy        241 SFREGLCRLLELTRRPIFITFEGQRRDPAHPALRLLRQHRHLVTLLLWRPGSVTPSSDFW 300
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SFREGLCRLLELTRRPIFITFEGQRRDPAHPALRLLRQHRHLVTLLLWRPGSVTPSSDFW 300

Qy        301 KEVQLALPRKVRYRPVEGDPQTQLQDDKDPMLILRGRVPEGRALDSEVDPDPEGDLGVRG 360
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 KEVQLALPRKVRYRPVEGDPQTQLQDDKDPMLILRGRVPEGRALDSEVDPDPEGDLGVRG 360

Qy        361 PVFGEPSAPPHTSGVSLGESRSSEVDVSDLGSRNYSARTDFYCLVSKDDM 410
                |||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 PVFGEPSAPPHTSGVSLGESRSSEVDVSDLGSRNYSARTDFYCLVSKDDM 410
```

RESULT 8
ADO20095
ID AD020095 standard; protein; 410 AA.
XX
AC AD020095;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human PRO polypeptide #502.

XX
KW Human; PRO; immune related disorder; systemic lupus erythematosus;
KW rheumatoid arthritis; osteoarthritis; juvenile chronic arthritis;
KW systemic sclerosis; Sjogren's syndrome; vasculitis; sarcoidosis;
KW autoimmune haemolytic anaemia; autoimmune thrombocytopenia; thyroiditis;
KW diabetes mellitus; renal disease; demyelinating disease;
KW central nervous system; peripheral nervous system;
KW demyelinating polyneuropathy; Guillain-Barre syndrome;
KW chronic inflammatory demyelinating polyneuropathy.
XX
OS Homo sapiens.
XX
PN WO2004043361-A2.
XX
PD 27-MAY-2004.
XX
PF 06-NOV-2003; 2003WO-US035268.
XX
PR 08-NOV-2002; 2002US-0425235P.
XX
PA (GETH) GENENTECH INC.
XX
PI Fong S, Dennis K, Clark H, Chiu H, Schoenfeld J, Williams PM;
PI Wood WI, Wu TD;
XX
DR WPI; 2004-420067/39.
DR N-PSDB; ADO20094.
XX
PT Novel PRO polypeptide e.g., PRO69614, PRO71106, or PRO86388 useful for
PT treating an immune related disorder such as systemic lupus erythematosus,
PT rheumatoid arthritis, osteoarthritis, juvenile chronic arthritis or
PT spondyloarthropathy.
XX
PS Claim 7; SEQ ID NO 1004; 1731pp; English.
XX
CC The invention relates to human PRO polypeptides and the polynucleotides
CC encoding them. The polypeptides and polynucleotides are useful for
CC treating and diagnosing immune related disorders in mammals. The immune
CC related disorders include systemic lupus erythematosus, rheumatoid
CC arthritis, osteoarthritis, juvenile chronic arthritis, systemic
CC sclerosis, Sjogren's syndrome, vasculitis, sarcoidosis, autoimmune
CC haemolytic anaemia, autoimmune thrombocytopenia, thyroiditis, diabetes
CC mellitus, immune-mediated renal disease, demyelinating diseases of the
CC central or peripheral nervous system, demyelinating polyneuropathy,
CC Guillain-Barre syndrome and chronic inflammatory demyelinating
CC polyneuropathy. This sequence represents a human PRO polypeptide of the
CC invention.
XX
SQ Sequence 410 AA;

Query Match 100.0%; Score 2147; DB 8; Length 410;
Best Local Similarity 100.0%; Pred. No. 1.2e-218;
Matches 410; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MPGVCDRAPDFLSPSEDQVLRPALGSSVALNCTAWVVSGPHCSLPSVQWLKDGLPLGIGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MPGVCDRAPDFLSPSEDQVLRPALGSSVALNCTAWVVSGPHCSLPSVQWLKDGLPLGIGG 60

Qy         61 HYSLHEYSWVKANLSEVLVSSVLGVNVTSTEVYGAFTCSIQNISFSSFTLQRAGPTSHVA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 HYSLHEYSWVKANLSEVLVSSVLGVNVTSTEVYGAFTCSIQNISFSSFTLQRAGPTSHVA 120

Qy        121 AVLASLLVLLALLLAALLYVKCRLNVLLWYQDAYGEVEINDGKLYDAYVSYSDCPEDRKF 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AVLASLLVLLALLLAALLYVKCRLNVLLWYQDAYGEVEINDGKLYDAYVSYSDCPEDRKF 180

Qy        181 VNFILKPQLERRRGYKLFLDDRDLLPRAEPSADLLVNLSRCRRLIVVLSDAFLSRAWCSH 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VNFILKPQLERRRGYKLFLDDRDLLPRAEPSADLLVNLSRCRRLIVVLSDAFLSRAWCSH 240

Qy        241 SFREGLCRLLELTRRPIFITFEGQRRDPAHPALRLLRQHRHLVTLLLWRPGSVTPSSDFW 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SFREGLCRLLELTRRPIFITFEGQRRDPAHPALRLLRQHRHLVTLLLWRPGSVTPSSDFW 300

Qy        301 KEVQLALPRKVRYRPVEGDPQTQLQDDKDPMLILRGRVPEGRALDSEVDPDPEGDLGVRG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 KEVQLALPRKVRYRPVEGDPQTQLQDDKDPMLILRGRVPEGRALDSEVDPDPEGDLGVRG 360

Qy        361 PVFGEPSAPPHTSGVSLGESRSSEVDVSDLGSRNYSARTDFYCLVSKDDM 410
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 PVFGEPSAPPHTSGVSLGESRSSEVDVSDLGSRNYSARTDFYCLVSKDDM 410
```

RESULT 9
ADO19087
ID ADO19087 standard; protein; 410 AA.
XX
AC ADO19087;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human PRO polypeptide #12.
XX
KW Human; PRO; immune related disorder; systemic lupus erythematosus;
KW rheumatoid arthritis; osteoarthritis; juvenile chronic arthritis;
KW systemic sclerosis; Sjogren's syndrome; vasculitis; sarcoidosis;
KW autoimmune haemolytic anaemia; autoimmune thrombocytopenia; thyroiditis;
KW diabetes mellitus; renal disease; demyelinating disease;
KW central nervous system; peripheral nervous system;
KW demyelinating polyneuropathy; Guillain-Barre syndrome;
KW chronic inflammatory demyelinating polyneuropathy.
XX
OS Homo sapiens.
XX
PN WO2004043361-A2.
XX
PD 27-MAY-2004.
XX
PF 06-NOV-2003; 2003WO-US035268.
XX
PR 08-NOV-2002; 2002US-0425235P.
XX
PA (GETH) GENENTECH INC.
XX
PI Fong S, Dennis K, Clark H, Chiu H, Schoenfeld J, Williams PM;
PI Wood WI, Wu TD;
XX
DR WPI; 2004-420067/39.
DR N-PSDB; ADO19086.
XX
PT Novel PRO polypeptide e.g., PRO69614, PRO71106, or PRO86388 useful for
PT treating an immune related disorder such as systemic lupus erythematosus,
PT rheumatoid arthritis, osteoarthritis, juvenile chronic arthritis or
PT spondyloarthropathy.
XX
PS Claim 7; SEQ ID NO 24; 1731pp; English.
XX
CC The invention relates to human PRO polypeptides and the polynucleotides
CC encoding them. The polypeptides and polynucleotides are useful for
CC treating and diagnosing immune related disorders in mammals. The immune
CC related disorders include systemic lupus erythematosus, rheumatoid
CC arthritis, osteoarthritis, juvenile chronic arthritis, systemic
CC sclerosis, Sjogren's syndrome, vasculitis, sarcoidosis, autoimmune
CC haemolytic anaemia, autoimmune thrombocytopenia, thyroiditis, diabetes
CC mellitus, immune-mediated renal disease, demyelinating diseases of the
CC central or peripheral nervous system, demyelinating polyneuropathy,
CC Guillain-Barre syndrome and chronic inflammatory demyelinating
CC polyneuropathy. This sequence represents a human PRO polypeptide of the
CC invention.
XX
SQ Sequence 410 AA;

Query Match 100.0%; Score 2147; DB 8; Length 410;
Best Local Similarity 100.0%; Pred. No. 1.2e-218;
Matches 410; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MPGVCDRAPDFLSPSEDQVLRPALGSSVALNCTAWVVSGPHCSLPSVQWLKDGLPLGIGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MPGVCDRAPDFLSPSEDQVLRPALGSSVALNCTAWVVSGPHCSLPSVQWLKDGLPLGIGG 60

Qy         61 HYSLHEYSWVKANLSEVLVSSVLGVNVTSTEVYGAFTCSIQNISFSSFTLQRAGPTSHVA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 HYSLHEYSWVKANLSEVLVSSVLGVNVTSTEVYGAFTCSIQNISFSSFTLQRAGPTSHVA 120

Qy        121 AVLASLLVLLALLLAALLYVKCRLNVLLWYQDAYGEVEINDGKLYDAYVSYSDCPEDRKF 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AVLASLLVLLALLLAALLYVKCRLNVLLWYQDAYGEVEINDGKLYDAYVSYSDCPEDRKF 180

Qy        181 VNFILKPQLERRRGYKLFLDDRDLLPRAEPSADLLVNLSRCRRLIVVLSDAFLSRAWCSH 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VNFILKPQLERRRGYKLFLDDRDLLPRAEPSADLLVNLSRCRRLIVVLSDAFLSRAWCSH 240

Qy        241 SFREGLCRLLELTRRPIFITFEGQRRDPAHPALRLLRQHRHLVTLLLWRPGSVTPSSDFW 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SFREGLCRLLELTRRPIFITFEGQRRDPAHPALRLLRQHRHLVTLLLWRPGSVTPSSDFW 300

Qy        301 KEVQLALPRKVRYRPVEGDPQTQLQDDKDPMLILRGRVPEGRALDSEVDPDPEGDLGVRG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 KEVQLALPRKVRYRPVEGDPQTQLQDDKDPMLILRGRVPEGRALDSEVDPDPEGDLGVRG 360

Qy        361 PVFGEPSAPPHTSGVSLGESRSSEVDVSDLGSRNYSARTDFYCLVSKDDM 410
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 PVFGEPSAPPHTSGVSLGESRSSEVDVSDLGSRNYSARTDFYCLVSKDDM 410
```

```
RESULT 10
ADP54078
ID   ADP54078 standard; protein; 410 AA.
XX
AC   ADP54078;
XX
DT   18-NOV-2004  (first entry)
XX
DE   Human PRO protein sequence SEQ ID NO:54.
XX
KW   human; PRO; immune related disease; inflammatory immune response;
KW   immune response stimulation; antiallergic; antianaemic; antiarthritic;
KW   antiasthmatic; antidiabetic; antiinflammatory; antipsoriatic;
KW   antirheumatic; antithyroid; CNS; dermatological; gastrointestinal;
KW   haemostatic; hepatotropic; immunostimulant; immunosuppressive; muscular;
KW   nephrotropic; neuroprotective; osteopathic; respiratory; vasotropic;
KW   virucide; gene therapy.
XX
OS   Homo sapiens.
XX
PN   WO2004039956-A2.
XX
PD   13-MAY-2004.
XX
PF   28-OCT-2003; 2003WO-US034381.
XX
PR   29-OCT-2002; 2002US-0422472P.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Aggarwal S,  Clark H,  Gurney AL,  Schoenfeld J,  Williams PM;
PI   Wood WI,  Wu TD;
XX
DR   WPI; 2004-376182/35.
DR   N-PSDB; ADP54077.
XX
PT   New PRO polynucleotides and polypeptides, useful in useful in diagnosing
PT   and treating an immune related disease, e.g. systemic lupus
PT   erythematosus, rheumatoid arthritis, diabetes mellitus or asthma and in
PT   stimulating an immune response.
XX
PS   Claim 1; SEQ ID NO 54; 3009pp; English.
XX
CC   The present invention describes an isolated PRO nucleic acid (I). Also
CC   described: (1) a vector comprising (I); (2) a host cell comprising the
CC   vector of (1); (3) a process for producing a PRO polypeptides; (4) an
CC   isolated PRO polypeptide; (5) a chimeric molecule comprising the
CC   polypeptide of (4) fused to a heterologous amino acid sequence; (6) an
CC   antibody which specifically binds to a polypeptide of (4); (7) a
CC   composition of matter comprising a polypeptide of (4), an agonist or
CC   antagonist of the polypeptide or an antibody that binds to the
CC   polypeptide in combination with a carrier; (8) an article of manufacture
CC   comprising a container, a label on the container and a composition of
CC   matter of (7); (9) a method of treating an immune related disease in a
CC   mammal; (10) a method for determining the presence of a PRO polypeptide
CC   in a sample suspected of having the polypeptide; (11) a method of
CC   diagnosing an immune related disease or an inflammatory immune response
CC   in mammal; (12) a method of identifying a compound that inhibits or
CC   mimics the activity of or expression of a gene encoding a PRO polypeptide
CC   ; and (13) a method of stimulating the immune response in a mammal. The
CC   PRO sequences have antiallergic, antianaemic, antiarthritic,
CC   antiasthmatic, antidiabetic, antiinflammatory, antipsoriatic,
CC   antirheumatic, antithyroid, CNS, dermatological, gastrointestinal,
CC   haemostatic, hepatotropic, immunostimulant, immunosuppressive, muscular,
CC   nephrotropic, neuroprotective, osteopathic, respiratory, vasotropic and
CC   virucide activities, and can be used in gene therapy. The nucleic acid
CC   (I) and the encoded polypeptides, compositions, kits and methods are
CC   useful in diagnosing and treating an immune related disease and in
CC   stimulating an immune response. The present sequence represents a human
CC   PRO protein from the present invention.
XX
SQ   Sequence 410 AA;
```

```
  Query Match             100.0%;  Score 2147;  DB 8;  Length 410;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-218;
  Matches  410;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MPGVCDRAPDFLSPSEDQVLRPALGSSVALNCTAWVVSGPHCSLPSVQWLKDGLPLGIGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MPGVCDRAPDFLSPSEDQVLRPALGSSVALNCTAWVVSGPHCSLPSVQWLKDGLPLGIGG 60

Qy         61 HYSLHEYSWVKANLSEVLVSSVLGVNVTSTEVYGAFTCSIQNISFSSFTLQRAGPTSHVA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 HYSLHEYSWVKANLSEVLVSSVLGVNVTSTEVYGAFTCSIQNISFSSFTLQRAGPTSHVA 120

Qy        121 AVLASLLVLLALLLAALLYVKCRLNVLLWYQDAYGEVEINDGKLYDAYVSYSDCPEDRKF 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AVLASLLVLLALLLAALLYVKCRLNVLLWYQDAYGEVEINDGKLYDAYVSYSDCPEDRKF 180

Qy        181 VNFILKPQLERRRGYKLFLDDRDLLPRAEPSADLLVNLSRCRRLIVVLSDAFLSRAWCSH 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VNFILKPQLERRRGYKLFLDDRDLLPRAEPSADLLVNLSRCRRLIVVLSDAFLSRAWCSH 240

Qy        241 SFREGLCRLLELTRRPIFITFEGQRRDPAHPALRLLRQHRHLVTLLLWRPGSVTPSSDFW 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SFREGLCRLLELTRRPIFITFEGQRRDPAHPALRLLRQHRHLVTLLLWRPGSVTPSSDFW 300

Qy        301 KEVQLALPRKVRYRPVEGDPQTQLQDDKDPMLILRGRVPEGRALDSEVDPDPEGDLGVRG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 KEVQLALPRKVRYRPVEGDPQTQLQDDKDPMLILRGRVPEGRALDSEVDPDPEGDLGVRG 360

Qy        361 PVFGEPSAPPHTSGVSLGESRSSEVDVSDLGSRNYSARTDFYCLVSKDDM 410
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 PVFGEPSAPPHTSGVSLGESRSSEVDVSDLGSRNYSARTDFYCLVSKDDM 410
```

```
RESULT 11
ADP22970
ID   ADP22970 standard; protein; 410 AA.
XX
AC   ADP22970;
XX
DT   18-NOV-2004  (first entry)
XX
DE   PRO polypeptide SEQ ID NO:64.
XX
KW   PRO; antiinflammatory; antiarthritic; antirheumatic; immunosuppressive;
```

KW osteopathic; antidiabetic; dermatological; antipsoriatic; antiallergic;
KW antiasthmatic; hepatotropic; respiratory; gene therapy; immune system.
XX
OS Unidentified.
XX
PN WO2004041170-A2.
XX
PD 21-MAY-2004.
XX
PF 30-OCT-2003; 2003WO-US034312.
XX
PR 01-NOV-2002; 2002US-0423394P.
XX
PA (GETH) GENENTECH INC.
XX
PI Clark H, Schoenfeld J, Van Lookeren M, Williams PM, Wood WI;
PI Wu TD;
XX
DR WPI; 2004-419628/39.
DR N-PSDB; ADP22969.
XX
PT New PRO polypeptides and polynucleotides, useful for treating e.g.
PT erythematosus, rheumatoid arthritis, diabetes mellitus, immune-mediated
PT renal disease, or demyelinating diseases of the central or peripheral
PT nervous system.
XX
PS Claim 7; SEQ ID NO 64; 2940pp; English.
XX
CC The invention relates to a novel isolated nucleic acid and the PRO
CC polypeptide encoded by it. A protein of the invention has
CC antiinflammatory, antiarthritic, antirheumatic, immunosuppressive,
CC osteopathic, antidiabetic, dermatological, antipsoriatic, antiallergic,
CC antiasthmatic, hepatotropic, and respiratory activity. A polynucleotide
CC of the invention may have a use in gene therapy. The PRO polypeptide, its
CC agonist, antagonist, or antibody that specifically binds to the
CC polypeptide is useful for treating an immune related disorder such as
CC systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis,
CC juvenile chronic arthritis, a spondyloarthropathy, systemic sclerosis, an
CC idiopathic inflammatory myopathy, Sjogren's syndrome, systemic
CC vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune
CC thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal
CC disease, a demyelinating disease of the central or peripheral nervous
CC system, idiopathic demyelinating polyneuropathy, Guillain-Barre syndrome,
CC a chronic inflammatory demyelinating polyneuropathy, a hepatobiliary
CC disease, infectious or autoimmune chronic active hepatitis, primary
CC biliary cirrhosis, granulomatous hepatitis, sclerosing cholangitis,
CC inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's
CC disease, an autoimmune or immune-mediated skin disease, a bullous skin
CC disease, erythema multiforme, contact dermatitis, psoriasis, an allergic
CC disease, asthma, allergic rhinitis, atopic dermatitis, food
CC hypersensitivity, urticaria, an immunologic disease of the lung,
CC eosinophilic pneumonia, idiopathic pulmonary fibrosis, hypersensitivity
CC pneumonitis, a transplantation associated disease, graft rejection or
CC graft-versus-host disease. The present sequence represents a PRO protein
CC of the invention.
XX
SQ Sequence 410 AA;

Query Match 100.0%; Score 2147; DB 8; Length 410;
Best Local Similarity 100.0%; Pred. No. 1.2e-218;
Matches 410; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MPGVCDRAPDFLSPSEDQVLRPALGSSVALNCTAWVVSGPHCSLPSVQWLKDGLPLGIGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MPGVCDRAPDFLSPSEDQVLRPALGSSVALNCTAWVVSGPHCSLPSVQWLKDGLPLGIGG 60

Qy         61 HYSLHEYSWVKANLSEVLVSSVLGVNVTSTEVYGAFTCSIQNISFSSFTLQRAGPTSHVA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 HYSLHEYSWVKANLSEVLVSSVLGVNVTSTEVYGAFTCSIQNISFSSFTLQRAGPTSHVA 120

Qy        121 AVLASLLVLLALLLAALLYVKCRLNVLLWYQDAYGEVEINDGKLYDAYVSYSDCPEDRKF 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AVLASLLVLLALLLAALLYVKCRLNVLLWYQDAYGEVEINDGKLYDAYVSYSDCPEDRKF 180

Qy        181 VNFILKPQLERRRGYKLFLDDRDLLPRAEPSADLLVNLSRCRRLIVVLSDAFLSRAWCSH 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VNFILKPQLERRRGYKLFLDDRDLLPRAEPSADLLVNLSRCRRLIVVLSDAFLSRAWCSH 240

Qy        241 SFREGLCRLLELTRRPIFITFEGQRRDPAHPALRLLRQHRHLVTLLLWRPGSVTPSSDFW 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SFREGLCRLLELTRRPIFITFEGQRRDPAHPALRLLRQHRHLVTLLLWRPGSVTPSSDFW 300

Qy        301 KEVQLALPRKVRYRPVEGDPQTQLQDDKDPMLILRGRVPEGRALDSEVDPDPEGDLGVRG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 KEVQLALPRKVRYRPVEGDPQTQLQDDKDPMLILRGRVPEGRALDSEVDPDPEGDLGVRG 360

Qy        361 PVFGEPSAPPHTSGVSLGESRSSEVDVSDLGSRNYSARTDFYCLVSKDDM 410
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 PVFGEPSAPPHTSGVSLGESRSSEVDVSDLGSRNYSARTDFYCLVSKDDM 410
```

RESULT 12
AAU86151
ID AAU86151 standard; protein; 504 AA.
XX
AC AAU86151;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human PRO342 polypeptide.
XX
KW Human; PRO; benign tumour; malignant tumour; lymphoid malignancy;
KW leukaemia; neuronal disorder; stromal disorder; blastocoelic disorder;
KW inflammatory disorder; immune disorder; angiogenic disorder; cytostatic;
KW neuroprotective.
XX
OS Homo sapiens.
XX
PN WO200153486-A1.
XX
PD 26-JUL-2001.
XX
PF 11-FEB-2000; 2000WO-US003565.
XX
PR 08-MAR-1999; 99WO-US005028.
PR 11-MAR-1999; 99US-0123972P.
PR 11-MAY-1999; 99US-0133459P.
PR 02-JUN-1999; 99WO-US012252.
PR 22-JUN-1999; 99US-0140650P.
PR 22-JUN-1999; 99US-0140653P.
PR 20-JUL-1999; 99US-0144758P.
PR 26-JUL-1999; 99US-0145698P.
PR 28-JUL-1999; 99US-0146222P.
PR 17-AUG-1999; 99US-0149395P.
PR 31-AUG-1999; 99US-0151689P.
PR 01-SEP-1999; 99WO-US020111.
PR 15-SEP-1999; 99WO-US021090.
PR 30-NOV-1999; 99WO-US028313.
PR 01-DEC-1999; 99WO-US028301.
PR 01-DEC-1999; 99WO-US028634.
PR 05-JAN-2000; 2000WO-US000219.
XX
PA (GETH) GENENTECH INC.
XX
PI Ashkenazi AJ, Goddard A, Godowski PJ, Gurney AL, Hillan KJ;
PI Marsters SA, Pan J, Pitti RM, Roy MA, Smith V, Stone DM;
PI Watanabe CK, Wood WI;
XX
DR WPI; 2002-205567/26.
DR N-PSDB; ABK40277.
XX
PT Thirty five nucleic acids encoding PRO polypeptides, useful for treating
PT benign or malignant tumors, leukemias and lymphoid malignancies,
PT inflammatory, angiogenic and immunologic disorders.
XX
PS Claim 61; Fig 48; 302pp; English.

```
XX
CC   The present invention relates to the isolation of novel human PRO
CC   polypeptides and the polynucleotide sequences encoding them. The PRO
CC   polypeptides, agonists, antagonists or anti-PRO antibodies are useful for
CC   treating benign or malignant tumours (e.g. renal, kidney, bladder,
CC   breast, etc), leukaemias and lymphoid malignancies, other disorders such
CC   as neuronal, glial, astrocytal, hypothalamic, glandular, macrophagal,
CC   stromal and blastocoelic disorders, inflammatory, immune and angiogenic
CC   disorders. The polynucleotide sequences are also useful in gene therapy.
CC   AAU86128-AAU86162 represent the human PRO polypeptides of the invention
XX
SQ   Sequence 504 AA;

  Query Match             87.4%;  Score 1876;  DB 5;  Length 504;
  Best Local Similarity   96.8%;  Pred. No. 1.1e-189;
  Matches  360;  Conservative    2;  Mismatches    2;  Indels    8;  Gaps    1;

Qy          1 MPGVCDRAPDFLSPSEDQVLRPALGSSVALNCTAWVVSGPHCSLPSVQWLKDGLPLGIGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MPGVCDRAPDFLSPSEDQVLRPALGSSVALNCTAWVVSGPHCSLPSVQWLKDGLPLGIGG 60

Qy         61 HYSLHEYSWVKANLSEVLVSSVLGVNVTSTEVYGAFTCSIQNISFSSFTLQRAGPTSHVA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 HYSLHEYSWVKANLSEVLVSSVLGVNVTSTEVYGAFTCSIQNISFSSFTLQRAGPTSHVA 120

Qy        121 AVLASLLVLLALLLAALLYVKCRLNVLLWYQDAYGEVEINDGKLYDAYVSYSDCPEDRKF 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AVLASLLVLLALLLAALLYVKCRLNVLLWYQDAYGEVEINDGKLYDAYVSYSDCPEDRKF 180

Qy        181 VNFILKPQLERRRGYKLFLDDRDLLPRAEPSADLLVNLSRCRRLIVVLSDAFLSRAWCSH 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VNFILKPQLERRRGYKLFLDDRDLLPRAEPSADLLVNLSRCRRLIVVLSDAFLSRAWCSH 240

Qy        241 SFREGLCRLLELTRRPIFITFEGQRRDPAHPALRLLRQHRHLVTLLLWRPGSVTPSSDFW 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SFREGLCRLLELTRRPIFITFEGQRRDPAHPALRLLRQHRHLVTLLLWRPGSVTPSSDFW 300

Qy        301 KEVQLALPRKVRYRPVEGDPQTQLQDDKDPMLILRGRVPEGRALDSEVDPDPEGDLGVRG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:
Db        301 KEVQLALPRKVRYRPVEGDPQTQLQDDKDPMLILRGRVPEGRALDSEVDPDPEGDLGM-- 358

Qy        361 PVFGEPSAPPHT 372
                    | ||:
Db        359 ------PAQPHS 364


RESULT 13
ADJ37329
ID   ADJ37329 standard; protein; 504 AA.
XX
AC   ADJ37329;
XX
DT   22-APR-2004  (first entry)
XX
DE   Human tumour therapy associated PRO342.
XX
KW   cytostatic; gene therapy; PRO; PRO197; PRO207; PRO226; PRO232; PRO243;
KW   PRO256; PRO269; PRO274; PRO304; PRO339; PRO1558; PRO779; PRO1185;
KW   PRO1245; PRO1759; PRO5775; PRO7133; PRO7168; PRO5725; PRO202; PRO206;
KW   PRO264; PRO313; PRO342; PRO542; PRO773; PRO861; PRO1216; PRO1686;
KW   PRO1800; PRO3562; PRO9850; PRO539; PRO4316; PRO4980; cancer; tumour;
KW   neoplastic cell growth; neoplastic cell proliferation; carcinoma;
KW   lymphoma; blastoma; sarcoma; leukaemia.
XX
OS   Homo sapiens.
XX
PN   US2003211096-A1.
XX
PD   13-NOV-2003.
XX
PF   02-AUG-2002; 2002US-00211858.
XX
PR   31-AUG-1999;   99US-0151689P.
PR   11-FEB-2000; 2000WO-US003565.
PR   09-AUG-2001; 2001US-00927796.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Ashkenazi AJ,  Goddard A,  Godowski PJ,  Gurney AL,  Hillan KJ;
PI   Marsters SA,  Pan J,  Pitti RM,  Roy MA,  Smith V,  Stone DM;
PI   Watanabe CK,  Wood WI;
XX
DR   WPI; 2003-901564/82.
DR   N-PSDB; ADJ37328.
XX
PT   New isolated PRO polypeptides, useful as targets for the diagnosis,
PT   prevention and treatment of cancers, e.g. lymphoma, blastoma, sarcoma or
PT   leukemia, and as predictors of the prognosis of tumor treatment.
XX
PS   Claim 61; SEQ ID NO 48; 307pp; English.
XX
CC   The invention describes an isolated PRO polypeptide. The PRO polypeptide:
CC   has at least 80% amino acid sequence identity to: (1) any one of 35 fully
CC   defined sequences of 104-954 amino acids (designated P1-P35) given in the
CC   specification, with or without its associated signal peptide; (2) an
CC   extracellular domain of any one of the polypeptides of P1-P35, with or
CC   without its associated signal peptide; or (3) an amino acid sequence
CC   encoded by the full-length coding sequence of the DNA deposited under
CC   ATCC accession number 209284, 209358, 203376, 209250, 209508, 209379,
CC   209397, 209786, 209482, 209490, 203312, 55820, 203096, 203155, 203465,
CC   PTA-255, PTA-618, PTA-545, PTA-256, 203538, 203661, 203835 or PTA-43; or
CC   scores at least 80% positives when compared to any one of the sequences
CC   of P1-P35. Specifically claimed are 35 PRO polypeptides, i.e. PRO197,
CC   PRO207, PRO226, PRO232, PRO243, PRO256, PRO269, PRO274, PRO304, PRO339,
CC   PRO1558, PRO779, PRO1185, PRO1245, PRO1759, PRO5775, PRO7133, PRO7168,
CC   PRO5725, PRO202, PRO206, PRO264, PRO313, PRO342, PRO542, PRO773, PRO861,
CC   PRO1216, PRO1686, PRO1800, PRO3562, PRO9850, PRO539, PRO4316 and PRO4980
CC   polypeptides. The PRO polypeptides are useful as targets for the
CC   diagnosis, prevention and treatment of cancers, and as predictors of the
CC   prognosis of tumour treatment. The nucleic acid molecules, antibodies and
CC   antagonists are useful for diagnosing and treating neoplastic cell growth
CC   and proliferation, e.g. carcinoma, lymphoma, blastoma, sarcoma or
CC   leukaemia. The antibodies may be used in screening assays for drug
CC   candidates. This is the amino acid sequence of a PRO protein useful in
CC   the treatment of cancers.
XX
SQ   Sequence 504 AA;

  Query Match             87.4%;  Score 1876;  DB 7;  Length 504;
  Best Local Similarity   96.8%;  Pred. No. 1.1e-189;
  Matches  360;  Conservative    2;  Mismatches    2;  Indels    8;  Gaps    1;

Qy          1 MPGVCDRAPDFLSPSEDQVLRPALGSSVALNCTAWVVSGPHCSLPSVQWLKDGLPLGIGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MPGVCDRAPDFLSPSEDQVLRPALGSSVALNCTAWVVSGPHCSLPSVQWLKDGLPLGIGG 60

Qy         61 HYSLHEYSWVKANLSEVLVSSVLGVNVTSTEVYGAFTCSIQNISFSSFTLQRAGPTSHVA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 HYSLHEYSWVKANLSEVLVSSVLGVNVTSTEVYGAFTCSIQNISFSSFTLQRAGPTSHVA 120

Qy        121 AVLASLLVLLALLLAALLYVKCRLNVLLWYQDAYGEVEINDGKLYDAYVSYSDCPEDRKF 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AVLASLLVLLALLLAALLYVKCRLNVLLWYQDAYGEVEINDGKLYDAYVSYSDCPEDRKF 180

Qy        181 VNFILKPQLERRRGYKLFLDDRDLLPRAEPSADLLVNLSRCRRLIVVLSDAFLSRAWCSH 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VNFILKPQLERRRGYKLFLDDRDLLPRAEPSADLLVNLSRCRRLIVVLSDAFLSRAWCSH 240

Qy        241 SFREGLCRLLELTRRPIFITFEGQRRDPAHPALRLLRQHRHLVTLLLWRPGSVTPSSDFW 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SFREGLCRLLELTRRPIFITFEGQRRDPAHPALRLLRQHRHLVTLLLWRPGSVTPSSDFW 300

Qy        301 KEVQLALPRKVRYRPVEGDPQTQLQDDKDPMLILRGRVPEGRALDSEVDPDPEGDLGVRG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:
Db        301 KEVQLALPRKVRYRPVEGDPQTQLQDDKDPMLILRGRVPEGRALDSEVDPDPEGDLGM-- 358
```

```
Qy        361 PVFGEPSAPPHT 372
                  | ||:
Db        359 ------PAQPHS 364
```

RESULT 14
ADG68253
ID ADG68253 standard; protein; 504 AA.
XX
AC ADG68253;
XX
DT 11-MAR-2004 (first entry)
XX
DE Human PRO polypeptide #24.
XX
KW neoplastic tumour; lung; colon; breast; prostate; rectal; cervical;
KW liver; gene therapy; Human.
XX
OS Homo sapiens.
XX
PN US2003170228-A1.
XX
PD 11-SEP-2003.
XX
PF 02-AUG-2002; 2002US-00210951.
XX
PR 31-AUG-1999; 99US-0151689P.
PR 11-FEB-2000; 2000WO-US003565.
PR 09-AUG-2001; 2001US-00927796.
XX
PA (GETH) GENENTECH INC.
XX
PI Ashkenazi AJ, Goddard A, Godowski PJ, Gurney AL, Hillan KJ;
PI Marsters SA, Pan J, Pitti RM, Roy MA, Smith V, Stone DM;
PI Watanabe CK, Wood WI;
XX
DR WPI; 2004-020650/02.
DR N-PSDB; ADG68252.
XX
PT New isolated antibodies binding PRO polypeptides, useful for diagnosing,
PT prognosticating and/or treating neoplastic tumors, such as lung, colon,
PT breast, prostate, rectal, cervical and liver tumors.
XX
PS Claim 61; SEQ ID NO 48; 308pp; English.
XX
CC The invention relates to an isolated antibody that binds to a PRO
CC polypeptide. The methods and compositions of the present invention are
CC useful for diagnosing, prognosticating and/or treating neoplastic
CC tumours, such as lung, colon, breast, prostate, rectal, cervical and
CC liver tumours. The PRO polypeptides are also useful as molecular weight
CC markers, or for chromosome identification. The PRO genes are useful as
CC hybridisation probes, or for screening libraries of Human cDNA, genomic
CC DNA or mRNA. The PRO genes may also be used in gene therapy, particularly
CC for replacing a defective gene. The present sequence is used in the
CC exemplification of the invention.
XX
SQ Sequence 504 AA;

```
  Query Match             87.4%;  Score 1876;  DB 8;  Length 504;
  Best Local Similarity   96.8%;  Pred. No. 1.1e-189;
  Matches  360;  Conservative    2;  Mismatches    2;  Indels    8;  Gaps    1;

Qy          1 MPGVCDRAPDFLSPSEDQVLRPALGSSVALNCTAWVVSGPHCSLPSVQWLKDGLPLGIGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MPGVCDRAPDFLSPSEDQVLRPALGSSVALNCTAWVVSGPHCSLPSVQWLKDGLPLGIGG 60

Qy         61 HYSLHEYSWVKANLSEVLVSSVLGVNVTSTEVYGAFTCSIQNISFSSFTLQRAGPTSHVA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 HYSLHEYSWVKANLSEVLVSSVLGVNVTSTEVYGAFTCSIQNISFSSFTLQRAGPTSHVA 120

Qy        121 AVLASLLVLLALLLAALLYVKCRLNVLLWYQDAYGEVEINDGKLYDAYVSYSDCPEDRKF 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AVLASLLVLLALLLAALLYVKCRLNVLLWYQDAYGEVEINDGKLYDAYVSYSDCPEDRKF 180

Qy        181 VNFILKPQLERRRGYKLFLDDRDLLPRAEPSADLLVNLSRCRRLIVVLSDAFLSRAWCSH 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VNFILKPQLERRRGYKLFLDDRDLLPRAEPSADLLVNLSRCRRLIVVLSDAFLSRAWCSH 240

Qy        241 SFREGLCRLLELTRRPIFITFEGQRRDPAHPALRLLRQHRHLVTLLLWRPGSVTPSSDFW 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SFREGLCRLLELTRRPIFITFEGQRRDPAHPALRLLRQHRHLVTLLLWRPGSVTPSSDFW 300

Qy        301 KEVQLALPRKVRYRPVEGDPQTQLQDDKDPMLILRGRVPEGRALDSEVDPDPEGDLGVRG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:
Db        301 KEVQLALPRKVRYRPVEGDPQTQLQDDKDPMLILRGRVPEGRALDSEVDPDPEGDLGM-- 358

Qy        361 PVFGEPSAPPHT 372
                  | ||:
Db        359 ------PAQPHS 364
```

RESULT 15
AAU17408
ID AAU17408 standard; protein; 407 AA.
XX
AC AAU17408;
XX
DT 07-NOV-2001 (first entry)
XX
DE Novel signal transduction pathway protein, Seq ID 973.
XX
KW Neuroprotective; cytostatic; dermatological; immunosuppressive; tumour;
KW antiinflammatory; anti-HIV; antibacterial; antiinflammatory; cancer;
KW immune system disorder; rheumatoid arthritis; inflammatory condition;
KW organ transplant rejection; infection; hepatitis C; blood disorder;
KW sickle cell anaemia; hyperproliferative disorder; Gaucher's disease;
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KW chromosomal abnormality; Down syndrome; ischaemia; renal disorder;
KW cardiovascular; respiratory; wound healing; endocrine; Addison's disease;
KW reproductive system; gastrointestinal; liver disorder; AIDS;
KW acquired immune deficiency syndrome.
XX
OS Homo sapiens.
XX
PN WO200154733-A1.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001312.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.

PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-465460/50.
DR N-PSDB; AAS27325.
XX
PT Novel polypeptides useful for diagnosing, treating, preventing and/or
PT prognosing disorders related to the proteins, including cancers, immune
PT disorders and neuronal disorders.
XX
PS Claim 1; SEQ ID NO 973; 880pp; English.
XX
CC The invention relates to novel isolated polypeptides (I), and
CC polynucleotides (II). (I), (II) and the antibody to (I) are useful for
CC diagnosing, preventing and treating diseases including immune system
CC disorders (e.g. congenital and acquired immunodeficiencies, autoimmune
CC disorders (e.g rheumatoid arthritis), inflammatory conditions, organ
CC transplant rejections and graft versus host disease, infectious diseases
CC (e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and
CC other blood-related disorders (sickle cell anaemia), myeloproliferative
CC disorders, primary haematopoietic disorders, hyperproliferative disorders
CC (e.g. Gaucher's disease and cancer), neurodegenerative disorders (e.g.
CC Alzheimer's disease, Parkinson's disease), chromosomal abnormalities
CC (Down syndrome), ischaemic injury (e.g. stroke), renal disorders (e.g.
CC glomerulonephritis), cardiovascular disorders (e.g. arrhythmia),
CC respiratory disorders, dermatological disorders, in wound healing,
CC epithelial cell proliferation, endocrine disorders (e.g. Addison's
CC disease), reproductive system disorders, gastrointestinal disorder
CC (inflammatory disorders), liver disorders (cirrhosis), as stimulators of
CC B-cell responsiveness to pathogens, activators of T-cells, to induce
CC higher affinity antibodies, and as a means to induce tumour proliferation
CC in pathologies e.g. acquired immune deficiency syndrome (AIDS). AAU17059-
CC AAU17683 represent novel signal transduction pathway protein, amino acid
CC sequences of the invention

XX

Query Match 82.0%; Score 1761; DB 4; Length 407;
Best Local Similarity 100.0%; Pred. No. 1.3e-177;
Matches 341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         70 VKANLSEVLVSSVLGVNVTSTEVYGAFTCSIQNISFSSFTLQRAGPTSHVAAVLASLLVL 129
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 VKANLSEVLVSSVLGVNVTSTEVYGAFTCSIQNISFSSFTLQRAGPTSHVAAVLASLLVL 126

Qy        130 LALLLAALLYVKCRLNVLLWYQDAYGEVEINDGKLYDAYVSYSDCPEDRKFVNFILKPQL 189
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        127 LALLLAALLYVKCRLNVLLWYQDAYGEVEINDGKLYDAYVSYSDCPEDRKFVNFILKPQL 186

Qy        190 ERRRGYKLFLDDRDLLPRAEPSADLLVNLSRCRRLIVVLSDAFLSRAWCSHSFREGLCRL 249
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 ERRRGYKLFLDDRDLLPRAEPSADLLVNLSRCRRLIVVLSDAFLSRAWCSHSFREGLCRL 246

Qy        250 LELTRRPIFITFEGQRRDPAHPALRLLRQHRHLVTLLLWRPGSVTPSSDFWKEVQLALPR 309
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 LELTRRPIFITFEGQRRDPAHPALRLLRQHRHLVTLLLWRPGSVTPSSDFWKEVQLALPR 306

Qy        310 KVRYRPVEGDPQTQLQDDKDPMLILRGRVPEGRALDSEVDPDPEGDLGVRGPVFGEPSAP 369
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 KVRYRPVEGDPQTQLQDDKDPMLILRGRVPEGRALDSEVDPDPEGDLGVRGPVFGEPSAP 366

Qy        370 PHTSGVSLGESRSSEVDVSDLGSRNYSARTDFYCLVSKDDM 410
              |||||||||||||||||||||||||||||||||||||||||
Db        367 PHTSGVSLGESRSSEVDVSDLGSRNYSARTDFYCLVSKDDM 407
```

Search completed: August 1, 2005, 13:53:00
Job time : 168 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: August 1, 2005, 13:34:31 ; Search time 41 Seconds
(without alignments)
962.168 Million cell updates/sec

Title: US-09-598-443-2
Perfect score: 2147
Sequence: 1 MPGVCDRAPDFLSPSEDQVL.....GSRNYSARTDFYCLVSKDDM 410

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	215.5	10.0	570	2	A57535	intrleukin 1 recep
2	209.5	9.8	569	2	A36187	interleukin-1 rece
3	204	9.5	562	2	G02426	interleukin-1 rece
4	190.5	8.9	576	2	A32604	interleukin-1 rece
5	189.5	8.8	567	2	S29498	lymphocyte antigen
6	184	8.6	555	2	JQ1526	interleukin-1 rece
7	176	8.2	1097	2	A29943	Toll protein precu
8	168.5	7.8	590	2	I56526	interleukin 1 rece
9	163	7.6	247	2	S42633	Fit-1M protein - r
10	144.5	6.7	1385	2	T13887	tlr protein - frui
11	139.5	6.5	1389	2	T13852	gene wheeler prote
12	128.5	6.0	169	2	I51903	type I interleukin
13	123	5.7	799	2	S18209	fibroblast growth
14	115	5.4	243	2	S11226	MyD88 protein - mo
15	114.5	5.3	786	2	T08664	Toll protein-like
16	113.5	5.3	1323	2	PN0568	connectin 3B - chi
17	111	5.2	26926	1	I38344	titin, cardiac mus
18	109.5	5.1	822	2	S19947	fibroblast growth
19	109.5	5.1	822	2	B49151	fibroblast growth
20	103	4.8	650	1	JC1450	fibroblast growth
21	102.5	4.8	818	2	JC4058	fibroblast growth
22	100.5	4.7	629	2	E70589	protoporphyrin IX
23	99	4.6	379	2	G96754	hypothetical prote
24	98	4.6	363	2	B96754	Similar to part of
25	97	4.5	176	2	E96754	Similar to downy m
26	97	4.5	829	2	JC4583	fibroblast growth
27	96.5	4.5	380	2	C96754	Similar to part of
28	96.5	4.5	802	1	TVHUF4	fibroblast growth
29	95	4.4	815	2	T41490	hypothetical prote
30	94.5	4.4	411	2	T35102	probable transmemb
31	94.5	4.4	1024	2	G71434	probable limonene
32	94	4.4	365	2	JC7780	coxsackie- and ade
33	94	4.4	1896	2	T08851	Down syndrome cell
34	93.5	4.4	275	2	D96754	Similar to part of
35	93.5	4.4	1179	2	T04584	TMV resistance pro
36	93.5	4.4	1344	2	T14316	rig-1 protein - mo
37	93.5	4.4	7962	2	I38346	elastic titin - hu
38	93	4.3	764	2	A49448	irregular chiasm C
39	92.5	4.3	371	2	F96754	Similar to NL27 [i
40	92.5	4.3	2783	2	T34416	hypothetical prote
41	92	4.3	871	1	I48696	protein-tyrosine k
42	92	4.3	881	1	I48697	protein-tyrosine k
43	92	4.3	1694	2	S50065	sialoadhesin - mou
44	92	4.3	2149	2	C96695	ribulose bisphosph
45	91.5	4.3	357	1	TVHUL2	transforming prote

ALIGNMENTS

RESULT 1
A57535
intrleukin 1 receptor accessory protein precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 08-Feb-1996 #sequence_revision 08-Feb-1996 #text_change 09-Jul-2004
C;Accession: A57535
R;Greenfeder, S.A.; Nunes, P.; Kwee, L.; Labow, M.; Chizzonite, R.A.; Ju, G.
J. Biol. Chem. 270, 13757-13765, 1995
A;Title: Molecular cloning and characterization o a second subunit of the interleukin 1
A;Reference number: A57535; MUID:95293970; PMID:7775431
A;Accession: A57535
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-570 <GRE>
A;Cross-references: UNIPROT:Q61730; GB:X85999; NID:g887520; PIDN:CAA59991.1; PID:g88752
C;Superfamily: interleukin-1 receptor type I
C;Keywords: transmembrane protein

Query Match 10.0%; Score 215.5; DB 2; Length 570;
Best Local Similarity 25.6%; Pred. No. 4.9e-10;
Matches 85; Conservative 61; Mismatches 153; Indels 33; Gaps 10;

```
Qy          9 PDFLSPSEDQVLRPALGSSVALNCTAW--VVSGPHCSLPSVQWLKDG-LPLGIGGHYSLH 65
              |   ||::  |      |  : : |  :   :   |      | |  ||  |  :    :::
Db        243 PQIYSPNDRVVYEKEPGEELVIPCKVYFSFIMDSH---NEVWWTIDGKKPDDVTVDITIN 299

Qy         66 EYSWVKANLSEVLVSSVLGVNVTSTEVYGAFTCSIQN----ISFSSFTLQRAGP---TSH 118
              |     :  |      :    ||  ::   : | :|       ::   |:  |   |
Db        300 ESVSYSSTEDETRTQILSIKKVTPEDLRRNYVCHARNTKGEAEQAAKVKQKVIPPRYTVE 359

Qy        119 VAAVLASLLVLLALLLAALLYVKCRLNVLLWYQDAYGEVE-INDGKLYDAYVSYSDCPED 177
              :|    : : |: :|:  ::|    | ::|:|:  :|  | | ||| || ||||:   |:
Db        360 LACGFGATVFLVVVLI--VVYHVYWLEMVLFYRAHFGTDETILDGKEYDIYVSYARNVEE 417

Qy        178 RKFVNFILKPQLERRRGYKLFLDDRDLLPRAEPSADLLVNLSRCRRLIVVLSDAFLSRAW 237
              :||   |:  ||   |||| : ||| ||    : : |  : : |||:||||  ::
Db        418 EEFVLLTLRGVLENEFGYKLCIFDRDSLPGGIVTDETLSFIQKSRRLLVVLSPNYV---- 473

Qy        238 CSHSFREGLCRLLELTRRPIFITFEG-------QRRDPAHPALRLLRQHRHLVTLLLWR- 289
                :|   ||||      :   |       | :      :: |:: : ::|:: |:
Db        474 -----LQGTQALLELKAGLENMASRGNINVILVQYKAVKDMKVKELKRAKTVLTVIKWKG 528

Qy        290 PGSVTPSSDFWKEVQLALPRKVRYRPVEGDPQ 321
              | |   |||::|:|:| |   |    | |
Db        529 EKSKYPQGRFWKQLQVAMPVKKSPRWSSNDKQ 560
```

RESULT 2
A36187
interleukin-1 receptor type I precursor - human
C;Species: Homo sapiens (man)
C;Date: 14-Dec-1990 #sequence_revision 14-Dec-1990 #text_change 09-Jul-2004

C;Accession: A36187; S06928
R;Sims, J.E.; Acres, R.B.; Grubin, C.E.; McMahan, C.J.; Wignall, J.M.; March, C.J.; Dowe
Proc. Natl. Acad. Sci. U.S.A. 86, 8946-8950, 1989
A;Title: Cloning the interleukin 1 receptor from human T cells.
A;Reference number: A36187; MUID:90046906; PMID:2530587
A;Accession: A36187
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-569 <SIM>
A;Cross-references: UNIPROT:P14778; GB:M20658
R;Chua, A.O.; Gubler, U.
Nucleic Acids Res. 17, 10114, 1989
A;Title: Sequence of the cDNA for the human fibroblast type interleukin-1 receptor.
A;Reference number: S06928; MUID:90098789; PMID:2532321
A;Accession: S06928
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-569 <CHU>
A;Cross-references: EMBL:X16896; NID:g33800; PIDN:CAA34773.1; PID:g33801
C;Genetics:
A;Gene: GDB:IL1R1; IL1RA; D2S1473; IL1R
A;Cross-references: GDB:125254; OMIM:147810
A;Map position: 2q12-2q12
C;Superfamily: interleukin-1 receptor type I
C;Keywords: cytokine receptor; transmembrane protein

Query Match 9.8%; Score 209.5; DB 2; Length 569;
Best Local Similarity 23.7%; Pred. No. 1.6e-09;
Matches 85; Conservative 62; Mismatches 126; Indels 85; Gaps 16;

```
Qy          7 RAPDFLSPSEDQVLRPA------------LGSSVALNCTAWVVSGPHCSLPSVQW----L 50
              |  :|::  |::  ||               ||| : | |    |:|    :   :|    :
Db        211 RVIEFITLEENKPTRPVIVSPANETMEVDLGSQIQLICN---VTGQLSDIAYWKWNGSVI 267

Qy         51 KDGLPLGIGGHYSLHEYSWVKANLSEVLVSSVLGVNVTSTEVY-GAFTCSIQNI-SFSSF 108
               :  |:    :||:       ||      : :|| ::   :  |   |||  :|     :
Db        268 DEDDPVLGEDYYSVEN----PANKRRSTLITVLNISEIESRFYKHPFTCFAKNTHGIDAA 323

Qy        109 TLQRAGPTS----HVAAVLASLLVLLALLLAALLYVKCRLNVLLWYQDA---YGEVEIND 161
              :|   | :    |:  :  :| |:  :: :  :|   :::::|||:|:   :  :: :|
Db        324 YIQLIYPVTNFQKHMIGICVTLTVI--IVCSVFIYKIFKIDIVLWYRDSCYDFLPIKASD 381

Qy        162 GKLYDAYVSY---------SDCPEDRKFVNFILKPQLERRRGYKLFLDDRDLLLPRAEPSA 212
              || ||||: |         |||     ||  :|   ||:: |||||:  ||     :
Db        382 GKTYDAYILYPKTVGEGSTSDCD---IFVFKVLPEVLEKQCGYKLFIYGRDDYVGEDIVE 438

Qy        213 DLLVNLSRCRRLIVVLSDAFLSRAWCSHSFREGLCR------------LLEL------TR 254
              :  |: : ||||::|       :|   |  | :               ||||       :
Db        439 VINENVKKSRRLIIILVRETSGFSWLGGSSEEQIAMYNALVQDGIKVVLLELEKIQDYEK 498

Qy        255 RPIFITFEGQRRDPAHPALRLLRQHRHLVTLLLW----RPGSVTPSSDFWKEVQLALP 308
              |  | |  |:    | |:|             |     |  :  : ||| |:  :|
Db        499 MPESIKFIKQK----HGAIR------------WSGDFTQGPQSAKTRFWKNVRYHMP 539
```

RESULT 3
G02426
interleukin-1 receptor-related protein - human
C;Species: Homo sapiens (man)
C;Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 29-Sep-1999
C;Accession: G02426
R;Lovenberg, T.W.; Crowe, P.D.; Liu, C.; Chalmers, D.T.; Liu, X.
submitted to the EMBL Data Library, February 1996
A;Reference number: H01239
A;Accession: G02426
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-562 <LOV>
A;Cross-references: EMBL:U49065; NID:g1236078; PIDN:AAB53237.1; PID:g1236079
C;Superfamily: interleukin-1 receptor type I

Query Match 9.5%; Score 204; DB 2; Length 562;
Best Local Similarity 28.5%; Pred. No. 4.5e-09;
Matches 81; Conservative 45; Mismatches 110; Indels 48; Gaps 15;

```
Qy         57 GIGGHYSLHEYSWVKANLSEVLVSSVLGVNVTSTEV----YG-AFTCSIQNISFSSFTLQ 111
              |:  | |  |:              ::  ||:|  ||    ||  | |    :| :   ||
Db        283 GVETHVSFREH------------NLYTVNITFLEVKMEDYGLPFMCH-AGVSTAYIILQ 328

Qy        112 RAGPTSHVAAVLASLLVLLALLLAAL-LYVKCRLNVLLWYQDAYGEVE-INDGKLYDAYV 169
                 |    | ::  |: |:|: :: : :|   :::::|||: |:   | | |||||||||
Db        329 LPAPDFR-AYLIGGLIALVAVAVSVVYIYNIFKIDIVLWYRSAFHSTETIVDGKLYDAYV 387

Qy        170 SYSDCPEDRK---------FVNFILKPQLERRRGYKLFLDDRDLLPRAEPSADLLVNLSR 220
              |   |: |           |  ||   |||: |||||:  ||  |    :  :  |:
Db        388 LY---PKPHKESQRHAVDALVLNILPEVLERQCGYKLFIFGRDEFPGQAVANVIDENVKL 444

Qy        221 CRRLIVVLSDAFLSRAWCSHSFREGLC---RLLELTRRPIFITFEGQRRDPAHP-ALRLL 276
              ||||||::    |      :  | :     |::   : | |  |          | ::: :
Db        445 CRRLIVIVVPESLGFGLLKNLSEEQIAVYSALIQDGMKVILIELEKIEDYTVMPESIQYI 504

Qy        277 RQHRHLVTLLLWRPGSVTPSSD-----FWKEVQLALPRKVRYRP 315
              :|    :    | | |  |       ||| |:  :| : | ||
Db        505 KQKHGAIR---WH-GDFTEQSQCMKTKFWKTVRYHMPPR-RCRP 543
```

RESULT 4
A32604
interleukin-1 receptor precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 09-Jul-2004
C;Accession: A32604
R;Sims, J.E.; March, C.J.; Cosman, D.; Widmer, M.B.; MacDonald, H.R.; McMahan, C.J.; Gr
S.K.
Science 241, 585-589, 1988
A;Title: cDNA expression cloning of the IL-1 receptor, a member of the immunoglobulin s
A;Reference number: A32604; MUID:88290679; PMID:2969618
A;Accession: A32604
A;Molecule type: mRNA
A;Residues: 1-576 <SIM>
A;Cross-references: UNIPROT:P13504; GB:M20658; GB:M29752; NID:g198300; PIDN:AAA39279.1;
C;Superfamily: interleukin-1 receptor type I
C;Keywords: cytokine receptor; glycoprotein; phosphoprotein; transmembrane protein
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-576/Product: interleukin-1 receptor #status predicted <INT>
F;339-359/Domain: transmembrane #status predicted <TMM>
F;46-99,145-199,251-315/Disulfide bonds: #status predicted
F;63,103,174,236,252,266,300/Binding site: carbohydrate (Asn) (covalent) #status predic

Query Match 8.9%; Score 190.5; DB 2; Length 576;
Best Local Similarity 24.0%; Pred. No. 6.3e-08;
Matches 86; Conservative 71; Mismatches 151; Indels 51; Gaps 18;

```
Qy          6 DRAPDFLSPSEDQVLRPALGSSVALNCTAWVVSGPHCSLPSVQWLKDGLPLGIGGHYSLH 65
              || |  |||  :: :    || : | |    |:|    |  | |  :|  :     :
Db        227 DR-PVILSP-RNETIEADPGSMIQLICN---VTGQFSDL--VYWKWNGSEIEWNDPFLAE 279

Qy         66 EYSWVK--ANLSEVLVSSVLGVNVTSTEVYG-AFTCSIQNIS-FSSFTLQRAGPTSHVAA 121
              :| :|:  :   :  : : | ::   :: |   | | ::| : | |  :|   |
Db        280 DYQFVEHPSTKRKYTLITTLNISEVKSQFYRYPFICVVKNTNIFESAHVQLIYPVPDFKN 339

Qy        122 VLASLLVLL--ALLLAALLYVKCRLNVLLWYQDA---YGEVEINDGKLYDAYV------- 169
                |    ::|   ::    :|   :::::|||:|:   :   : :||| ||||:
Db        340 YLIGGPIILTATIVCCVCIYKVFKVDIVLWYRDSCSGFLPSKASDGKTYDAYILYPKTLG 399

Qy        170 --SYSDCPEDRKFVNFILKPQLERRRGYKLFLDDRDLLPRAEPSADLLVNLSRCRRLIVV 227
                |:||      ||  :|   || : |||||:  ||      :       |: : ||||::
Db        400 EGSFSDLD---TFVFKLLPEVLEGQPGYKLFIYGRDDYVGEDTIEVTNENVKKSRRLIII 456

Qy        228 LSDAFLSRAWCSHSFREGLCRLLELTRRPIFITF----EGQRRDPAHPALRLLRQHRHLV 283
              |       :|   |  | :     | :  | |       : |  :    ::: ::| :| |
Db        457 LVRDMGGFSWLGQSSEEQIAIYNALIQEGIKIVLLELEKIQDYEKMPDSIQFIKQ-KHGV 515

Qy        284 TLLLW------RPGSVTPSSDFWKEVQLALPRKVR-----YRPVEGDPQTQLQDDKDPM 331
              : |      || |    : ||| ::  :| : |      :| :  ||   ::| |: :
Db        516 --ICWSGDFQERPQSA--KTRFWKNLRYQMPAQRRSPLSKHRLLTLDP---VRDTKEKL 567
```

```
RESULT 5
S29498
lymphocyte antigen Ly84 precursor - mouse
N;Alternate names: 38.5K T1 glycoprotein; ST2L protein
C;Species: Mus musculus (house mouse)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C;Accession: S29498; A33541; S17657; S07054
R;Yanagisawa, K.; Takagi, T.; Tsukamoto, T.; Tetsuka, T.; Tominaga, S.
FEBS Lett. 318, 83-87, 1993
A;Title: Presence of a novel primary response gene ST2L, encoding a product highly simil
A;Reference number: S29498; MUID:93170492; PMID:7916701
A;Accession: S29498
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-567 <YAN>
A;Cross-references: UNIPROT:P14719; EMBL:D13695; NID:g286100; PIDN:BAA02854.1; PID:g2861
R;Klemenz, R.; Hoffmann, S.; Werenskiold, A.K.
Proc. Natl. Acad. Sci. U.S.A. 86, 5708-5712, 1989
A;Title: Serum- and oncoprotein-mediated induction of a gene with sequence similarity to
A;Reference number: A33541; MUID:89345536; PMID:2527364
A;Accession: A33541
A;Molecule type: mRNA
A;Residues: 1-191,'V',193-328,'SKECPSHIA' <KLE>
A;Cross-references: GB:M24843; NID:g201103; PIDN:AAA40160.1; PID:g201104
R;Tominaga, S.I.; Jenkins, N.A.; Gilbert, D.J.; Copeland, N.G.; Tetsuka, T.
Biochim. Biophys. Acta 1090, 1-8, 1991
A;Title: Molecular cloning of the murine ST2 gene. Characterization and chromosomal mapp
A;Reference number: S17657; MUID:91355215; PMID:1832015
A;Accession: S17657
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-191,'V',193-328,'SKECPSHIA' <TOM>
A;Cross-references: EMBL:X60184; NID:g54200; PIDN:CAA42742.1; PID:g54201
R;Tominaga, S.I.
FEBS Lett. 258, 301-304, 1989
A;Title: A putative protein of a growth specific cDNA from BALB/c-3T3 cells is highly si
A;Reference number: S07054; MUID:90092495; PMID:2532153
A;Accession: S07054
A;Molecule type: mRNA
A;Residues: 1-328,'SKECPSHIA' <TO2>
A;Cross-references: EMBL:Y07519; NID:g55517; PIDN:CAA68812.1; PID:g55518
A;Note: it is uncertain whether Met-1, Met-7 or Met-19 is the initiator
C;Genetics:
A;Gene: ST2
A;Map position: 1
A;Introns: 27/1; 97/2; 155/3; 210/1; 233/1; 280/2
C;Superfamily: interleukin-1 receptor type I
C;Keywords: glycoprotein
F;1-26/Domain: signal sequence #status predicted <SIG>
F;27-337/Product: ST2 protein #status predicted <MAT>
F;60,101,107,146,176,194,225,259,278/Binding site: carbohydrate (Asn) (covalent) #status

  Query Match             8.8%;  Score 189.5;  DB 2;  Length 567;
  Best Local Similarity   24.8%;  Pred. No. 7.4e-08;
  Matches   82;  Conservative   59;  Mismatches  153;  Indels   37;  Gaps   14;

Qy          9 PDFLSPSEDQVLRPALGSSVALNCTAWVVSGPHCSLPSVQWLKDGLPLGIGGHYSLHEYS 68
              |   :|  :  :   :|   :: |:|    | |  |  | |  :   :|  |   : |
Db        217 PVITNPPYNHTMEVEIGKPASIACSACFGKGSH-FLADVLWQINKTVVGNPGEARIQEEE 275

Qy         69 WVKANLSEVL--VSSVLGV-NVTSTEV---YGAFTCSIQNISFSSFTLQRAGPTSH--VA 120
                : |  :  ::||| :  ||  ::   |     ::  :   :  |:|  |  |  :
Db        276 GRNESSSNDMDCLTSVLRITGVTEKDLSLEYDCLALNLHGMIRHTIRLRRKQPIDHRSIY 335

Qy        121 AVLASLLVLLALLLAALLYVKC-RLNVLLWYQDAYGEVEI-NDGKLYDAYVSY------- 171
               ::|   :||  :   :: :|   : | |:::|     :  |||||||||: |
Db        336 YIVAGCSLLLMFINVLVIVLKVFWIEVALFWRDIVTPYKTRNDGKLYDAYIIYPRVFRGS 395

Qy        172 -SDCPEDRKFVNFILKPQLERRRGYKLFLDDRDLLPRAEPSADLLVNLSRCRRLIVVLSD 230
              :       ||:  |   || : |||| :  |||||  : :  :  ::   || : ||:
Db        396 AAGTHSVEYFVHHTLPDVLENKCGYKLCIYGRDLLPGQDAATVVESSIQNSRRQVFVLAP 455

Qy        231 AFLSRAWCSHSFREGL-CRLLELTRRPIFITFEGQRRDPAHPALRL----LRQH-RHLVT 284
              :     ::      |   |::   : | |  |      |   | ||    |:   :|||
Db        456 HMMHSKEFAYEQEIALHSALIQNNSKVILIEME-----PLGEASRLQVGDLQDSLQHLVK 510

Qy        285 L---LLWRPGSVTP----SSDFWKEVQLALP 308
              :   : ||   |       || ||| |:  :|
Db        511 IQGTIKWREDHVADKQSLSSKFWKHVRYQMP 541


RESULT 6
JQ1526
interleukin-1 receptor I precursor - chicken
C;Species: Gallus gallus (chicken)
C;Date: 17-Aug-1992 #sequence_revision 17-Aug-1992 #text_change 09-Jul-2004
C;Accession: JQ1526
R;Guida, S.; Heguy, A.; Melli, M.
Gene 111, 239-243, 1992
A;Title: The chicken IL-1 receptor: differential evolution of the cytoplasmic and extra
A;Reference number: JQ1526; MUID:92175529; PMID:1531799
A;Accession: JQ1526
A;Molecule type: mRNA
A;Residues: 1-555 <GUI>
A;Cross-references: UNIPROT:Q90874; GB:M81846; NID:g212206; PIDN:AAA48924.1; PID:g21220
A;Experimental source: embryo
C;Superfamily: interleukin-1 receptor type I
C;Keywords: cytokine receptor; transmembrane protein
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-555/Product: interleukin-1 receptor I #status predicted <INT>
F;339-359/Domain: transmembrane #status predicted <TRA>

  Query Match             8.6%;  Score 184;  DB 2;  Length 555;
  Best Local Similarity   25.0%;  Pred. No. 2.1e-07;
  Matches   90;  Conservative   65;  Mismatches  101;  Indels  104;  Gaps   21;

Qy          9 PDFLSPSEDQVLRPALGSSVALNCTAWVVSGPHCSLP----------------SVQWLKD 52
              |:|: |: :  :   ||| | : |   | || :  ||                  |: ::
Db        230 PEFIYPN-NNTIEVELGSHVVMECN--VSSGVYGLLPYWQVNDEDVDSFDSTYREQFYEE 286

Qy         53 GLPLGIGGHYSLHEYSWVKANLSEVLVSSVLGVNVTSTEVYGAFTCSI-QNISFSSFTLQ 111
              |:| ||          |  | |:||| :              |  |   |  :  |:|: ::
Db        287 GMPHGIA-------VSGTKFNISEVKLKDY---------AYKFFCHFIYDSQEFTSY-IK 329

Qy        112 RAGPTSHVAAVL----ASLLVLLALLLAALLYVKCRLNVLLWYQDAYGEV---EINDGKL 164
                  ::   |     ||: || |:|  ::|   :::::|||: :   :   :::|||:
Db        330 LEHPVQNIRGYLIGGGISLIFLLFLIL--IVYKIFKIDIVLWYRSSCHPLLGKKVSDGKI 387

Qy        165 YDAYVSYSDCPEDRK--------FVNFILKPQLERRRGYKLFLDDRDLLPRAEPSADLL- 215
              ||||| |   |::|:         |   ||   |||: || ||:  |: |   |   |:
Db        388 YDAYVLY---PKNRESCLYSSDIFALKILPEVLERQCGYNLFIFGRNDL-AGEAVIDVTD 443

Qy        216 VNLSRCRRLIVVL-------------SDAFLSRAWCSHSFREGL-CRLLELTR------R 255
              : : ||:|::|              |:  |  |  :   ::|:   |:|| :
Db        444 EKIHQSRRVIIILVPEPSCYGILEDASEKHL--AVYNALIQDGIKIILIELEKIEDYANM 501

Qy        256 PIFITFEGQRRDPAHPALRLL-----RQHRHLVTLLLWRPGSVTPSSDFWKEVQLALPRK 310
              |  | :  |:    : |:|       | |               : |: |||:|:  :| :
Db        502 PESIKYVKQK----YGAIRWTGDFSERSH-------------SASTRFWKKVRYHMPSR 543


RESULT 7
A29943
Toll protein precursor - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 19-Nov-1988 #sequence_revision 19-Nov-1988 #text_change 09-Jul-2004
C;Accession: A29943
R;Hashimoto, C.; Hudson, K.L.; Anderson, K.V.
Cell 52, 269-279, 1988
A;Title: The Toll gene of Drosophila, required for dorsal-ventral embryonic polarity, a
A;Reference number: A29943; MUID:88135760; PMID:2449285
A;Accession: A29943
A;Molecule type: DNA
A;Residues: 1-1097 <HAS>
```

```
A;Cross-references: UNIPROT:P08953; GB:M19969; GB:J02682; NID:g158640; PIDN:AAA28941.1;
C;Genetics:
A;Gene: FlyBase:Tl
A;Cross-references: FlyBase:FBgn0003717
C;Keywords: transmembrane protein
F;1-17/Domain: signal sequence #status predicted <SIG>
F;18-1097/Product: Toll protein #status predicted <MAT>

  Query Match             8.2%;  Score 176;  DB 2;  Length 1097;
  Best Local Similarity   25.7%;  Pred. No. 2.3e-06;
  Matches   70;  Conservative   47;  Mismatches   89;  Indels   66;  Gaps   13;

Qy        121 AVLASLLVLLALLLAALLY-----VKCRL---NVLLWYQDAYGEVEINDGKLYDAYVSYS 172
              ||: :|  |||   ||| |     :|  |   |:|||:     | :::  | :||::|||
Db        810 AVVIALTGLLAGFTAALYYKFQTEIKIWLYAHNLLLWF---VTEEDLDKDKKFDAFISYS 866

Qy        173 DCPEDRKFVNFILKPQLER-RRGYKLFLDDRDLLPRAEPSADLLVNLSRCRRLIVVLSDA 231
                :|: |:   | ||||   : ::| : :|| |       ::: :::  || |:|||
Db        867 H--KDQSFIEDYLVPQLEHGPQKFQLCVHERDWLVGGHIPENIMRSVADSRRTIIVLSQN 924

Qy        232 FLSRAWCSHSFR--------EGLCRLLELTRRPIFITFEGQRRDPAHPALRLLRQHRHLV 283
              |:   |    ||        ||  |:: :    |    : :: |        |: :  :
Db        925 FIKSEWARLEFRAAHRSALNEGRSRIIVIIYSDI---GDVEKLDEE------LKAYLKMN 975

Qy        284 TLLLWRPGSVTPSSD--FWKEVQLALPRKVRYRPVEGDPQTQLQDDKDPMLILRGRVPEG 341
              | | |          |  || ::: ||| :   |||                  | :  |
Db        976 TYLKW--------GDPWFWDKLRFALPHR---RPV------------------GNIGNG 1005

Qy        342 RALDSEVDPDPEGDLGVRGPVFGEPSAPPHTS 373
                : : :    :  | :  |    |  || |:
Db       1006 ALIKTALKGSTDDKLELIKP---SPVTPPLTT 1034


RESULT 8
I56526
interleukin 1 receptor type I - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 29-Sep-1999
C;Accession: I56526
R;Hart, R.P.; Liu, C.; Shadiack, A.M.; McCormack, R.J.; Jonakait, G.M.
J. Neuroimmunol. 44, 49-56, 1993
A;Title: An mRNA homologous to interleukin-1 receptor type I is expressed in cultured ra
A;Reference number: I56526; MUID:93266794; PMID:7684399
A;Accession: I56526
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-590 <RES>
A;Cross-references: GB:M95578; NID:g451305; PIDN:AAA16196.1; PID:g451306
C;Superfamily: interleukin-1 receptor type I
C;Keywords: cytokine receptor

  Query Match             7.8%;  Score 168.5;  DB 2;  Length 590;
  Best Local Similarity   24.8%;  Pred. No. 4.5e-06;
  Matches   88;  Conservative   56;  Mismatches  128;  Indels   83;  Gaps   21;

Qy          6 DRAPDFLSPSEDQVLRPALGSSVALNCTAWVVSGPHCSLPSVQWLKDGLPLGIGGHYSLH 65
              || |  :||  :: :    ||:: | |    |:|    |  | |  :|  :
Db        241 DR-PVIMSP-RNETMEADPGSTIQLICN---VTGQFTDL--VYWKWNGSEIEWDDPILAE 293

Qy         66 EYSWVK---ANLSEVLVSSVLGVNVTSTEVYG-AFTCSIQNISFSSFTLQRAGPTSHVAA 121
              :| :::   |     |::: | |:   :: |   | | ::|             |:||
Db        294 DYQFLEHPSAKRKYTLITT-LNVSEVKSQFYRYPFICFVKNTHILE--------TAHVRL 344

Qy        122 VLA------SLLVLLALLLAALLYVKC-----RLNVLLWYQDAYGEV---EINDGKLYDA 167
              |          |:   |:  |  ::  |     :::::|||:|:  :    : :||: |||
Db        345 VYPVPDFKNYLIGGFAIFTATAVFCACIYKVFKVDIVLWYRDSCSDFLPRKASDGRTYDA 404

Qy        168 YVSYSDCPEDRKFV---NFILK--PQ-LERRRGYKLFLDDRDLLPRAEPSADLLVNLSRC 221
              || |     :  |     |: |  |: || : |||||:  ||     :         |: |
Db        405 YVLYPKTYGEGSFAYLDTFVFKLLPEVLEGQFGYKLFICGRDDYVGEDTIEVTNENVKRS 464

Qy        222 RRLIVVLSDAFLSRAWCSHS-----------FREGL-CRLLEL------TRRPIFITFEG 263
              ||||::|     | :    |             |||:   ||||       : |  | |
Db        465 RRLIIILVRDMGSFSCLGQSSEEQIAIYDALIREGIKIILLELEKIQDYEKMPESIQFIK 524

Qy        264 QRRDPAHPALRLLRQHRHLVTLLLW------RPGSVTPSSDFWKEVQLALPRKVR 312
              |:    | |             : |      || |    : ||| ::  :| : |
Db        525 QK----HGA-------------ICWSGDFKERPQSA--KTRFWKNLRYQMPAQRR 560


RESULT 9
S42633
Fit-1M protein - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 25-Dec-1994 #sequence_revision 10-Nov-1995 #text_change 03-Feb-2003
C;Accession: S42633
R;Bergers, G.; Reikerstorfer, A.; Braselmann, S.; Graninger, P.; Busslinger, M.
EMBO J. 13, 1176-1188, 1994
A;Title: Alternative promoter usage of the Fos-responsive gene Fit-1 generates mRNA isoi
A;Reference number: S42632; MUID:94178260; PMID:8131748
A;Accession: S42633
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-247 <BER>

  Query Match             7.6%;  Score 163;  DB 2;  Length 247;
  Best Local Similarity   28.1%;  Pred. No. 4.4e-06;
  Matches   65;  Conservative   39;  Mismatches   95;  Indels   32;  Gaps   11;

Qy        110 LQRAGPTSHVAA--VLASLLVLLALLLAALLYVKC-RLNVLLWYQDAYGEVEI-NDGKLY 165
              |:|  |  | :   ::|   :||  :   :: :|   : | |:::|     :  ||||||
Db          3 LRRKQPIDHQSTYYIVAGCSLLLMFINVLVIVLKVFWIEVALFWRDIMAPYKTQNDGKLY 62

Qy        166 DAYVSY--------SDCPEDRKFVNFILKPQLERRRGYKLFLDDRDLLPRAEPSADLLVN 217
              |||: |         :        ||:: |   || : |||| :  |||||  : :  :  :
Db         63 DAYIIYPRVFRGSAAGTGSVEYFVHYTLPQVLENKCGYKLQIYGRDLLPGQDAATVVESS 122

Qy        218 LSRCRRLIVVLSDAFLSRAWCSHSFREGL-CRLLELTRRPIFITFEGQRRDPAHPALRL- 275
              :   || : ||:  :     ::      |   |::   : | |  |     |   | ||
Db        123 IQNSRRQVFVLAPHMMHSKEFAYEQEIALHSALIQNNSKVILIEME-----PMGEASRLQ 177

Qy        276 ---LRQH-RHLVTL---LLWRPGSVTP----SSDFWKEVQLALPRKVRYRP 315
                 |:   :||| :   : ||   |      || ||| |:  :|  |  ||
Db        178 LGDLQDSLQHLVKMQGTIKWREDHVADKQSLSSKFWKHVRYQMP--VPKRP 226


RESULT 10
T13887
tlr protein - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C;Accession: T13887
R;Chiang, C.; Beachy, P.A.
Mech. Dev. 47, 225-239, 1994
A;Title: Expression of a novel Toll-like gene spans the parasegment boundary and contri
A;Reference number: Z17805; MUID:95151581; PMID:7848870
A;Accession: T13887
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1385 <CHI>
A;Cross-references: UNIPROT:Q24591; EMBL:S76155; NID:g913247; PID:g913248; PIDN:AAB3338
C;Genetics:
A;Cross-references: FlyBase:FBgn0004364
A;Note: tlr

  Query Match             6.7%;  Score 144.5;  DB 2;  Length 1385;
  Best Local Similarity   26.1%;  Pred. No. 0.0014;
  Matches   68;  Conservative   44;  Mismatches  104;  Indels   45;  Gaps   12;

Qy         99 SIQNISFSSFTLQRAGPTSHVAAVLASLLVLLALLLAALLYVKCRLNVLLWYQDAYG--- 155
              |  ||| |       ||   :   :||::|||: | :  ::    | :| :|    ||
Db        979 SASNISSSQ---DLAG--GYRLPLLAAVLVLIFLDVVLIIVFVFRESVRMWLFAHYGVRV 1033

Qy        156 -EVEIND-GKLYDAYVSYSDCPEDRKFVNFILKPQLER-RRGYKLFLDDRDLLPRAEPSA 212
               |   | |||||| : :|:  :| :||   :  :||  |  ::| :  ||| |:|
Db       1034 CEPRFEDAGKLYDAIILHSE--KDYEFVCRNIAAELEHGRPPFRLCIQQRDLPPQAS-HL 1090
```

```
Qy        213 DLLVNLSRCRRLIVVLSDAFLSRAWCSHSFREGLCRLLE-LTRRPIFI---TFEGQRRDP 268
              |:       |::|:||:   |: |    ||       |  | :: : |   :   : |
Db       1091 QLVEGARASRKIILVLTRNLLATEWNRIEFRNAFHESLRGLAQKLVIIEETSVSAEAEDV 1150

Qy        269 AH--------PALRLLRQHRHLVTLLLWRPGSVTPSSDFWKEVQLALPRKV--RYRPVEG 318
              |         |: |||   |:                 ||:::: |:| ::  |
Db       1151 AELSPYLKSVPSNRLLTCDRY----------------FWEKLRYAIPIELSPRGNNYTL 1193

Qy        319 DPQTQLQDDKDPMLILRGRVP 339
              |   : :    | :| |   |
Db       1194 DHHERFKQPVSPGMIFRQAPP 1214


RESULT 11
T13852
gene wheeler protein - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C;Accession: T13852
R;Eldon, E.; Kooyer, S.; D'evelyn, D.; Duman, M.; Lawinger, P.; Botas, J.; Bellen, H.
Development 120, 885-899, 1994
A;Title: The Drosophila 18 wheeler is required for morphogenesis and has striking simila
A;Reference number: Z17796; MUID:95324375; PMID:7600965
A;Accession: T13852
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1389 <ELD>
A;Cross-references: UNIPROT:Q24591; EMBL:L23171; NID:g415682; PID:g1019104; PIDN:AAA7920
C;Genetics:
A;Gene: wheeler
A;Cross-references: FlyBase:FBgn0004364

  Query Match             6.5%;  Score 139.5;  DB 2;  Length 1389;
  Best Local Similarity  23.2%;  Pred. No. 0.0036;
  Matches   66;  Conservative   51;  Mismatches  105;  Indels   63;  Gaps   12;

Qy         75 SEVLVSSVLGVNVTSTEVYGAFTCSIQNISFSSFTLQRAGPTSHVAAVLASLLVLLALLL 134
              |::| :|   :: :| :: ||  |                            ::|||: |::
Db        973 SDLLDASASNIS-SSQDLAGAIGC----------------------PCWPAVLVLIFLVV 1009

Qy        135 AALLYVKCRLNVLLWYQDAYG----EVEIND-GKLYDAYVSYSDCPEDRKFVNFILKPQL 189
                ::    | :| :|    ||    |    | |||||| : :|:  :| :||   :  :|
Db       1010 VLIIVFVFRESVRMWLFAHYGVRVCEPRFEDAGKLYDAIILHSE--KDYEFVCRNIAAEL 1067

Qy        190 ER-RRGYKLFLDDRDLLPRAEPSADLLVNLSRCRRLIVVLSDAFLSRAWCSHSFREGLCR 248
              |  |  ::| :  ||| |:|     |:      |::|:||:   |:  |    ||
Db       1068 EHGRPPFRLCIQQRDLPPQAS-HLQLVEGARASRKIILVLTRNLLATEWNRIEFRNAFHE 1126

Qy        249 LLE-LTRRPIFI---TFEGQRRDPAH--------PALRLLRQHRHLVTLLLWRPGSVTPS 296
              |  | :: : |   :   :  | |         |: |||   |:
Db       1127 SLRGLAQKLVIIEETSVSAEAEDVAELSPYLKSVPSNRLLTCDRY--------------- 1171

Qy        297 SDFWKEVQLALPRKV--RYRPVEGDPQTQLQDDKDPMLILRGRVP 339
               ||:::: |:| ::  |      |   : :    | :| |   |
Db       1172 --FWEKLRYAIPIELSPRGNNYTLDHHERFKQPVSPGMIFRQAPP 1214


RESULT 12
I51903
type I interleukin-1 receptor - rat (fragment)
C;Species: Rattus sp. (rat)
C;Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 13-Sep-1998
C;Accession: I51903
R;Sutherland, D.B.; Varilek, G.W.; Neil, G.A.
Am. J. Physiol. 266, C1198-C1203, 1994
A;Title: Identification and characterization of the rat intestinal epithelial cell (IEC-
A;Reference number: I51903; MUID:94262728; PMID:8203483
A;Accession: I51903
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-169 <RES>
A;Cross-references: GB:S70238; NID:g546976
A;Experimental source: intestinal epithelial IEC-18 cells
C;Superfamily: interleukin-1 receptor type I
C;Keywords: cytokine receptor

  Query Match             6.0%;  Score 128.5;  DB 2;  Length 169;
  Best Local Similarity  28.1%;  Pred. No. 0.0021;
  Matches   50;  Conservative   29;  Mismatches   76;  Indels   23;  Gaps    6;

Qy        145 NVLLWYQDAYGEV---EINDGKLYDAYVSYSDCPEDRKFV---NFILK--PQ-LERRRGY 195
              :::|||:|:  :    : :||| ||||| |     :  |     |: |  |: || : ||
Db          1 DIVLWYRDSCSDFLPPKASDGKTYDAYVLYPKTYGEGSFAYLDTFVFKLLPEVLEGQFGY 60

Qy        196 KLFLDDRDLLPRAEPSADLLVNLSRCRRLIVVLSDAFLSRAWCSHSFREGLCRLLELTRR 255
              |||:  ||     :       |: | ||||::|     | :   |  | :     | |
Db         61 KLFICGRDDYVGEDTIEVTNENVKRSRRLIIILVRDMGSFSCLGQSSEEQIAIYDALIRE 120

Qy        256 PIFITF----EGQRRDPAHPALRLLRQHRHLVTLLLWRPGSVTPSSDFWKEVQLALPR 309
              | |      | |  :    ::: ::|          : |::  | || :  | |  |
Db        121 GIKIILLELEEIQDYEKMPESIQFIKQ----------KHGAICWSGDFKERPQSAKTR 168


RESULT 13
S18209
fibroblast growth factor receptor 4 precursor (clone 61) - mouse
N;Alternate names: tyrosine kinase Mpk-11
N;Contains: protein-tyrosine kinase (EC 2.7.1.112)
C;Species: Mus musculus (house mouse)
C;Date: 07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change 09-Jul-2004
C;Accession: S18209; S26751; S30497
R;Stark, K.L.; McMahon, J.A.; McMahon, A.P.
Development 113, 641-651, 1991
A;Title: FGFR-4, a new member of the fibroblast growth factor receptor family, expresse
A;Reference number: S18209; MUID:92146274; PMID:1723680
A;Accession: S18209
A;Molecule type: mRNA
A;Residues: 1-799 <STA1>
A;Cross-references: UNIPROT:Q03142; EMBL:X59927
R;Stark, K.L.
submitted to the EMBL Data Library, May 1991
A;Reference number: S26751
A;Accession: S26751
A;Molecule type: mRNA
A;Residues: 1-485,'QVVRAEAFG',486-799 <STA2>
A;Cross-references: EMBL:X59927; NID:g50968; PIDN:CAA42551.1; PID:g50969
R;Gilardi-Hebenstreit, P.; Nieto, M.A.; Frain, M.; Mattei, M.G.; Chestier, A.; Wilkinso
Oncogene 7, 2499-2506, 1992
A;Title: An Eph-related receptor protein tyrosine kinase gene segmentally expressed in
A;Reference number: S30496; MUID:93096484; PMID:1281307
A;Accession: S30497
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 611-667 <GIL>
A;Cross-references: EMBL:X57236; NID:g53187; PIDN:CAA40512.1; PID:g53188
C;Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; prote
C;Keywords: ATP; duplication; glycoprotein; growth factor receptor; phosphotransferase;
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-799/Product: fibroblast growth factor receptor 4 #status predicted <MAT>
F;19-366/Domain: extracellular #status predicted <EXT>
F;162-223/Domain: immunoglobulin homology <IMM>
F;367-387/Domain: transmembrane #status predicted <TMM>
F;388-799/Domain: intracellular #status predicted <INT>
F;462-747/Domain: protein kinase homology <KIN>
F;470-478/Region: protein kinase ATP-binding motif
F;54-98,169-221,268-330/Disulfide bonds: #status predicted
F;500,517,609/Active site: Lys, Glu, Asp #status predicted

  Query Match             5.7%;  Score 123;  DB 2;  Length 799;
  Best Local Similarity  22.1%;  Pred. No. 0.044;
  Matches  103;  Conservative   45;  Mismatches  136;  Indels  182;  Gaps   23;

Qy         24 LGSSVALNCTAWVVSGPHCSLPSVQWLK----DGLPLGIGGHYSLHEYSWVKANLSEVLV 79
              :|| | | |  :  : ||     :||||    :|   |  |   :         |:|||
Db        260 VGSDVELLCKVYSDAQPH-----IQWLKHVVINGSSFGADGFPYVQVLKTTDINISEV-- 312
```

```
Qy          80 SSVLGVNVTSTEVYGAFTCSIQN---ISFSS----------FTLQRAGP----TSHVAAV 122
               || :   | |  | :||    |   :|: |           |   | |     |  :  |
Db         313 -QVLYLRNVSAEDAGEYTCLAGNSIGLSYQSAWLTVLPEEDLTWTTATPEARYTDIILYV 371

Qy         123 LASLLVLLALLLAALLYVKCRLNVLLWYQDAYGEVEINDGKLYDAYVSYSDCPEDRKFVN 182
                 ||::|: |||| : : :          :  |    :   ||        |  |  |:|
Db         372 SGSLVLLVLLLLAGVYHRQV-------IRGHYSRQPVTIQKL-------SRFPLARQF-- 415

Qy         183 FILKPQLERR----------RGYKLFLDDRDLLPRAEPSADLLVNLSRCRRLIVVLSDAF 232
                  || |          || :|      ||          ||||           |
Db         416 -----SLESRSSGKSSLSLVRGVRLSSSGPPLLTG-------LVNL----------DLP 452

Qy         233 LSRAWCSHSFREGLCRLLELTR------RPI-------FITFEGQRRDPAHP-------- 271
               |   |              |  |      :|:       :  |    ||: |
Db         453 LDPLW------------EFPRDRLVLGKPLGEGCFGQVVRAEAFGMDPSRPDQTSTVAV 499

Qy         272 --------------------ALRLLRQHRHLVTLLLWRPGSVTPSSDFWKEVQLALPRKV 311
                                 ::|: :|:::: ||    |  |     :  |: |    :
Db         500 KMLKDNASDKDLADLVSEMEVMKLIGRHKNIINLL----GVCTQEGPLYVIVECAAKGNL 555

Qy         312 RYRPVEGDPQTQLQDDKDPMLILRGRVPEGRALDSEVDPDPEGDLGVRGPVFGEPSAPPH 371
               |                       || | | |       |  |:|     ||:    | |
Db         556 RE-------------------FLRARRPPG------PDLSPDGPRSSEGPL----SFPAL 586

Qy         372 TS-------GVSLGESRSSEVDVSDLGSRNYSARTDFYCLVSKDDM 410
                |       |:   |||  :    || :||          ||::||:
Db         587 VSCAYQVARGMQYLESR--KCIHRDLAARN--------VLVTEDDV 622
```

RESULT 14
S11226
MyD88 protein - mouse
C;Species: Mus musculus (house mouse)
C;Date: 18-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 05-Nov-1999
C;Accession: S11226; JC2079
R;Lord, K.A.; Hoffman-Liebermann, B.; Liebermann, D.A.
Oncogene 5, 1095-1097, 1990
A;Title: Nucleotide sequence and expression of a cDNA encoding MyD88, a novel myeloid di
A;Reference number: S11226; MUID:90326414; PMID:2374694
A;Accession: S11226
A;Molecule type: mRNA
A;Residues: 1-243 <LOR>
A;Cross-references: EMBL:X51397; NID:g53293; PIDN:CAA35762.1; PID:g53294
R;Hultmark, D.
Biochem. Biophys. Res. Commun. 199, 144-146, 1994
A;Title: Macrophage differentiation marker MyD88 is a member of the Toll/IL-1 receptor f
A;Reference number: JC2079; MUID:94168566; PMID:8123006
A;Contents: annotation

  Query Match 5.4%; Score 115; DB 2; Length 243;
  Best Local Similarity 29.2%; Pred. No. 0.046;
  Matches 33; Conservative 26; Mismatches 38; Indels 16; Gaps 5;

```
Qy         163 KLYDAYVSYSDCPEDRKFVNFILKPQLERRRGYKLFLDDRDLLPRA---EPSADLLVNLS 219
               :|:||:: |  || | :||  :::   :     || : |||:||       :::|:
Db         106 ELFDAFICY--CPNDIEFVQEMIRQLEQTDYRLKLCVSDRDVLPGTCVWSIASELIE--K 161

Qy         220 RCRRLIVVLSDAFLSRAWCSHSFREGLCRLLELT-----RRPIFITFEGQRRD 267
               ||||::||:|| :|    |    :      | |:     :| | | ::  ::|
Db         162 RCRRMVVVVSDDYLQSKECDFQTKFA----LSLSPGVQQKRLIPIKYKAMKKD 210
```

RESULT 15
T08664
Toll protein-like receptor DKFZp547I0610.1 - human
C;Species: Homo sapiens (man)
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C;Accession: T08664
R;Poustka, A.; Klein, M.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, May 1999
A;Reference number: Z16466
A;Accession: T08664
A;Molecule type: mRNA
A;Residues: 1-786 <POU>
A;Cross-references: UNIPROT:Q15399; EMBL:AL050262
A;Experimental source: fetal brain; clone DKFZp547I0610
C;Genetics:
A;Note: DKFZp547I0610.1

  Query Match 5.3%; Score 114.5; DB 2; Length 786;
  Best Local Similarity 22.8%; Pred. No. 0.22;
  Matches 49; Conservative 39; Mismatches 88; Indels 39; Gaps 9;

```
Qy         119 VAAVLASLLVLLALLLAALLYVKCRLNVLLWYQDAYGE-------------VEINDGKLY 165
               :  ::|::|||   : :  :|:      | ||     :                |:     :
Db         583 IVTIVATMLVLAVTVTSLCIYLD-----LPWYLRMVCQWTQTRRRARNIPLEELQRNLQF 637

Qy         166 DAYVSYSDCPEDRKFVNFILKPQLERRRGYKLFLDDRDLLPRAEPSADLLVNLSRCRRLI 225
               |::|||    |  :|   | | || : | :: | :|: :|      :::  : :  : |
Db         638 HAFISYSG--HDSFWVKNELLPNLE-KEGMQICLHERNFVPGKSIVENIITCIEKSYKSI 694

Qy         226 VVLSDAFLSRAWCSHS--------FREGLCRLLELTRRPIFITFEGQRRDP-AHPALRLL 276
               |||  |:   || :          | ||   |: :   ||       |   | ::  |: |
Db         695 FVLSPNFVQSEWCHYELYFAHHNLFHEGSNSLILILLEPI-----PQYSIPSSYHKLKSL 749

Qy         277 RQHRHLVTLLLWRPGSVTPSSDFWKEVQLALPRKV 311
                  |  | | | |   :    ||  :: |:  |:
Db         750 MARR---TYLEW-PKEKSKRGLFWANLRAAINIKL 780
```

Search completed: August 1, 2005, 13:50:08
Job time : 43 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: August 1, 2005, 13:33:56 ; Search time 172 Seconds
(without alignments)
1220.654 Million cell updates/sec

Title: US-09-598-443-2
Perfect score: 2147
Sequence: 1 MPGVCDRAPDFLSPSEDQVL.....GSRNYSARTDFYCLVSKDDM 410

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2147	100.0	410	2	Q9H733	Q9h733 homo sapien
2	2143	99.8	410	2	Q6IA17	Q6ia17 homo sapien
3	1876	87.4	504	2	Q6UXI3	Q6uxi3 homo sapien
4	1546.5	72.0	409	2	Q9JLZ8	Q9jlz8 mus musculu
5	267.5	12.5	686	1	IPL2_HUMAN	Q9np60 h x-linked
6	261	12.2	686	1	IPL2_MOUSE	Q9ers6 mus musculu
7	257	12.0	695	1	IPL1_MOUSE	P59823 mus musculu
8	256.5	11.9	696	1	IPL1_HUMAN	Q9nzn1 homo sapien
9	256.5	11.9	696	1	IPL1_PANTR	P60029 pan troglod
10	256.5	11.9	696	1	IPL1_RAT	P59824 rattus norv
11	255.5	11.9	696	1	IPL1_PONPY	Q7yq19 pongo pygma
12	254	11.8	537	1	IR18_MOUSE	Q61098 mus musculu
13	240.5	11.2	570	1	ILAP_MACMU	P59822 macaca mula
14	231.5	10.8	541	1	IR18_HUMAN	Q13478 homo sapien
15	228.5	10.6	570	1	ILAP_HUMAN	Q9nph3 homo sapien
16	224.5	10.5	539	2	Q8HXQ6	Q8hxq6 sus scrofa
17	220	10.2	582	2	Q8AXT5	Q8axt5 salmo salar
18	219	10.2	539	2	Q6YBY1	Q6yby1 felis silve
19	218.5	10.2	568	2	Q70YG5	Q70yg5 salmo salar
20	217	10.1	574	1	IRL2_MOUSE	Q9ers7 mus musculu
21	215.5	10.0	570	1	ILAP_MOUSE	Q61730 mus musculu
22	213.5	9.9	569	2	Q6NWP5	Q6nwp5 homo sapien
23	211	9.8	533	2	Q9DEE1	Q9dee1 gallus gall
24	210.5	9.8	570	1	ILAP_RAT	Q63621 rattus norv
25	210	9.8	533	2	Q9DEE5	Q9dee5 gallus gall
26	209.5	9.8	569	1	IL1R_HUMAN	P14778 homo sapien
27	207.5	9.7	560	2	Q6RI75	Q6ri75 macaca fasc
28	206.5	9.6	575	1	IRL2_HUMAN	Q9hb29 homo sapien
29	204.5	9.5	561	1	IRL2_RAT	Q62929 rattus norv
30	204.5	9.5	569	2	Q6NWP6	Q6nwp6 homo sapien
31	203.5	9.5	573	2	Q9TV71	Q9tv71 equus cabal
32	202	9.4	556	1	IRL1_HUMAN	Q01638 homo sapien
33	197	9.2	1058	2	Q76FN7	Q76fn7 tachypleus
34	196.5	9.2	1152	2	Q8WRE5	Q8wre5 anopheles g
35	196.5	9.2	1152	2	Q7PFC4	Q7pfc4 anopheles g
36	196.5	9.2	1346	2	Q9V477	Q9v477 drosophila
37	191.5	8.9	573	2	Q8C6P3	Q8c6p3 mus musculu
38	190.5	8.9	573	2	Q8C833	Q8c833 mus musculu
39	190.5	8.9	576	1	IL1R_MOUSE	P13504 mus musculu
40	189.5	8.8	567	1	IRL1_MOUSE	P14719 mus musculu
41	189	8.8	599	2	O95256	O95256 homo sapien
42	185	8.6	1356	2	Q8WRE2	Q8wre2 anopheles g
43	184	8.6	555	2	Q90874	Q90874 gallus gall
44	183.5	8.5	571	2	Q8QFN4	Q8qfn4 salmo salar
45	183	8.5	566	1	IRL1_RAT	Q62611 rattus norv

ALIGNMENTS

RESULT 1
Q9H733
ID Q9H733 PRELIMINARY; PRT; 410 AA.
AC Q9H733;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein FLJ21446 (SIGIRR protein).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Ota T., Suzuki Y., Obayashi M., Nishi T., Shibahara T.,
RA Tanaka T., Nakamura Y., Isogai T., Sugano S.;
RL Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AK025099; BAB15066.1; -.
DR EMBL; BC025953; AAH25953.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004888; F:transmembrane receptor activity; IEA.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR004075; IL1_receptor1.

```
DR   InterPro; IPR000157; TIR.
DR   Pfam; PF01582; TIR; 1.
DR   PRINTS; PR01537; INTRLKN1R1F.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS50104; TIR; 1.
SQ   SEQUENCE   410 AA;  45707 MW;  2A7A663D79567ED6 CRC64;

  Query Match           100.0%;  Score 2147;  DB 2;  Length 410;
  Best Local Similarity  100.0%;  Pred. No. 2.1e-173;
  Matches  410;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MPGVCDRAPDFLSPSEDQVLRPALGSSVALNCTAWVVSGPHCSLPSVQWLKDGLPLGIGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MPGVCDRAPDFLSPSEDQVLRPALGSSVALNCTAWVVSGPHCSLPSVQWLKDGLPLGIGG 60

Qy         61 HYSLHEYSWVKANLSEVLVSSVLGVNVTSTEVYGAFTCSIQNISFSSFTLQRAGPTSHVA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 HYSLHEYSWVKANLSEVLVSSVLGVNVTSTEVYGAFTCSIQNISFSSFTLQRAGPTSHVA 120

Qy        121 AVLASLLVLLALLLAALLYVKCRLNVLLWYQDAYGEVEINDGKLYDAYVSYSDCPEDRKF 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AVLASLLVLLALLLAALLYVKCRLNVLLWYQDAYGEVEINDGKLYDAYVSYSDCPEDRKF 180

Qy        181 VNFILKPQLERRRGYKLFLDDRDLLPRAEPSADLLVNLSRCRRLIVVLSDAFLSRAWCSH 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VNFILKPQLERRRGYKLFLDDRDLLPRAEPSADLLVNLSRCRRLIVVLSDAFLSRAWCSH 240

Qy        241 SFREGLCRLLELTRRPIFITFEGQRRDPAHPALRLLRQHRHLVTLLLWRPGSVTPSSDFW 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SFREGLCRLLELTRRPIFITFEGQRRDPAHPALRLLRQHRHLVTLLLWRPGSVTPSSDFW 300

Qy        301 KEVQLALPRKVRYRPVEGDPQTQLQDDKDPMLILRGRVPEGRALDSEVDPDPEGDLGVRG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 KEVQLALPRKVRYRPVEGDPQTQLQDDKDPMLILRGRVPEGRALDSEVDPDPEGDLGVRG 360

Qy        361 PVFGEPSAPPHTSGVSLGESRSSEVDVSDLGSRNYSARTDFYCLVSKDDM 410
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 PVFGEPSAPPHTSGVSLGESRSSEVDVSDLGSRNYSARTDFYCLVSKDDM 410


RESULT 2
Q6IA17
ID   Q6IA17      PRELIMINARY;      PRT;   410 AA.
AC   Q6IA17;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   SIGIRR protein.
GN   Name=SIGIRR;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Ebert L., Schick M., Neubert P., Schatten R., Henze S., Korn B.;
RL   Submitted (JUN-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; CR457338; CAG33619.1; -.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0004888; F:transmembrane receptor activity; IEA.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR004075; IL1_receptor1.
DR   InterPro; IPR000157; TIR.
DR   Pfam; PF01582; TIR; 1.
DR   PRINTS; PR01537; INTRLKN1R1F.
DR   SMART; SM00255; TIR; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS50104; TIR; 1.
SQ   SEQUENCE   410 AA;  45689 MW;  2CBA663D79567ED6 CRC64;

  Query Match           99.8%;  Score 2143;  DB 2;  Length 410;
  Best Local Similarity  99.8%;  Pred. No. 4.5e-173;
  Matches  409;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MPGVCDRAPDFLSPSEDQVLRPALGSSVALNCTAWVVSGPHCSLPSVQWLKDGLPLGIGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MPGVCDRAPDFLSPSEDQVLRPALGSSVALNCTAWVVSGPHCSLPSVQWLKDGLPLGIGG 60

Qy         61 HYSLHEYSWVKANLSEVLVSSVLGVNVTSTEVYGAFTCSIQNISFSSFTLQRAGPTSHVA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 HYSLHEYSWVKANLSEVLVSSVLGVNVTSTEVYGAFTCSIQNISFSSFTLQRAGPTSHVA 120

Qy        121 AVLASLLVLLALLLAALLYVKCRLNVLLWYQDAYGEVEINDGKLYDAYVSYSDCPEDRKF 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AVLASLLVLLALLLAALLYVKCRLNVLLWYQDAYGEVEINDGKLYDAYVSYSDCPEDRKF 180

Qy        181 VNFILKPQLERRRGYKLFLDDRDLLPRAEPSADLLVNLSRCRRLIVVLSDAFLSRAWCSH 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VNFILKPQLERRRGYKLFLDDRDLLPRAEPSADLLVNLSRCRRLIVVLSDAFLSRAWCSH 240

Qy        241 SFREGLCRLLELTRRPIFITFEGQRRDPAHPALRLLRQHRHLVTLLLWRPGSVTPSSDFW 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SFREGLCRLLELTRRPIFITFEGQRRDPAHPALRLLRQHRHLVTLLLWRPGSVTPSSDFW 300

Qy        301 KEVQLALPRKVRYRPVEGDPQTQLQDDKDPMLILRGRVPEGRALDSEVDPDPEGDLGVRG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 KEVQLALPRKVRYRPVEGDPQTQLQDDKDPMLILRGRVPEGRALDSEVDPDPEGDLGVRG 360

Qy        361 PVFGEPSAPPHTSGVSLGESRSSEVDVSDLGSRNYSARTDFYCLVSKDDM 410
              |||||||||||||||||||||||||||||||||||||||||||||||||||:
Db        361 PVFGEPSAPPHTSGVSLGESRSSEVDVSDLGSRNYSARTDFYCLVSKDDI 410


RESULT 3
Q6UXI3
ID   Q6UXI3      PRELIMINARY;      PRT;   504 AA.
AC   Q6UXI3;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   SIGIRR.
GN   ORFNames=UNQ301;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22887296; PubMed=12975309; DOI=10.1101/gr.1293003;
RA   Clark H.F., Gurney A.L., Abaya E., Baker K., Baldwin D., Brush J.,
RA   Chen J., Chow B., Chui C., Crowley C., Currell B., Deuel B., Dowd P.,
RA   Eaton D., Foster J., Grimaldi C., Gu Q., Hass P.E., Heldens S.,
RA   Huang A., Kim H.S., Klimowski L., Jin Y., Johnson S., Lee J.,
RA   Lewis L., Liao D., Mark M., Robbie E., Sanchez C., Schoenfeld J.,
RA   Seshagiri S., Simmons L., Singh J., Smith V., Stinson J., Vagts A.,
RA   Vandlen R., Watanabe C., Wieand D., Woods K., Xie M.H., Yansura D.,
RA   Yi S., Yu G., Yuan J., Zhang M., Zhang Z., Goddard A., Wood W.I.,
RA   Godowski P.;
RT   "The secreted protein discovery initiative (SPDI), a large-scale
RT   effort to identify novel human secreted and transmembrane proteins: a
RT   bioinformatics assessment.";
RL   Genome Res. 13:2265-2270(2003).
DR   EMBL; AY358342; AAQ88708.1; -.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0004888; F:transmembrane receptor activity; IEA.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR004075; IL1_receptor1.
DR   InterPro; IPR000157; TIR.
DR   Pfam; PF01582; TIR; 1.
DR   PRINTS; PR01537; INTRLKN1R1F.
DR   SMART; SM00255; TIR; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS50104; TIR; 1.
SQ   SEQUENCE   504 AA;  55299 MW;  B3D4A5A60AF19D28 CRC64;

  Query Match           87.4%;  Score 1876;  DB 2;  Length 504;
```

```
  Best Local Similarity   96.8%;  Pred. No. 2.6e-150;
  Matches  360;  Conservative    2;  Mismatches    2;  Indels    8;  Gaps    1;

Qy          1 MPGVCDRAPDFLSPSEDQVLRPALGSSVALNCTAWVVSGPHCSLPSVQWLKDGLPLGIGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MPGVCDRAPDFLSPSEDQVLRPALGSSVALNCTAWVVSGPHCSLPSVQWLKDGLPLGIGG 60

Qy         61 HYSLHEYSWVKANLSEVLVSSVLGVNVTSTEVYGAFTCSIQNISFSSFTLQRAGPTSHVA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 HYSLHEYSWVKANLSEVLVSSVLGVNVTSTEVYGAFTCSIQNISFSSFTLQRAGPTSHVA 120

Qy        121 AVLASLLVLLALLLAALLYVKCRLNVLLWYQDAYGEVEINDGKLYDAYVSYSDCPEDRKF 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AVLASLLVLLALLLAALLYVKCRLNVLLWYQDAYGEVEINDGKLYDAYVSYSDCPEDRKF 180

Qy        181 VNFILKPQLERRRGYKLFLDDRDLLPRAEPSADLLVNLSRCRRLIVVLSDAFLSRAWCSH 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 VNFILKPQLERRRGYKLFLDDRDLLPRAEPSADLLVNLSRCRRLIVVLSDAFLSRAWCSH 240

Qy        241 SFREGLCRLLELTRRPIFITFEGQRRDPAHPALRLLRQHRHLVTLLLWRPGSVTPSSDFW 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SFREGLCRLLELTRRPIFITFEGQRRDPAHPALRLLRQHRHLVTLLLWRPGSVTPSSDFW 300

Qy        301 KEVQLALPRKVRYRPVEGDPQTQLQDDKDPMLILRGRVPEGRALDSEVDPDPEGDLGVRG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||:
Db        301 KEVQLALPRKVRYRPVEGDPQTQLQDDKDPMLILRGRVPEGRALDSEVDPDPEGDLGM-- 358

Qy        361 PVFGEPSAPPHT 372
                  |.||:
Db        359 ------PAQPHS 364


RESULT 4
Q9JLZ8
ID   Q9JLZ8      PRELIMINARY;      PRT;   409 AA.
AC   Q9JLZ8;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Toll/interleukin-1 receptor 8.
GN   Name=AI256711; Synonyms=TIR8;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   PubMed=14993616; DOI=10.1073/pnas.0308680101;
RA   Garlanda C., Riva F., Polentarutti N., Buracchi C., Sironi M.,
RA   De Bortoli M., Muzio M., Bergottini R., Scanziani E., Vecchi A.,
RA   Hirsch E., Mantovani A.;
RT   "Intestinal inflammation in mice deficient in Tir8, an inhibitory
RT   member of the IL-1 receptor family.";
RL   Proc. Natl. Acad. Sci. U.S.A. 101:3522-3526(2004).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Penton-Rol G., Polentarutti N., Muzio M., Mantovani A.;
RL   Submitted (DEC-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF113795; AAF26200.1; -.
DR   MGD; MGI:1344402; AI256711.
DR   GO; GO:0016021; C:integral to membrane; TAS.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR004075; IL1_receptor1.
DR   InterPro; IPR000157; TIR.
DR   Pfam; PF00047; ig; 1.
DR   Pfam; PF01582; TIR; 1.
DR   PRINTS; PR01537; INTRLKN1R1F.
DR   SMART; SM00409; IG; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS50104; TIR; 1.
KW   Receptor.
SQ   SEQUENCE   409 AA;  45707 MW;  DC4AB99A72356A13 CRC64;


  Query Match           72.0%;  Score 1546.5;  DB 2;  Length 409;
  Best Local Similarity   72.7%;  Pred. No. 1.7e-122;
  Matches  303;  Conservative   32;  Mismatches   67;  Indels   15;  Gaps    4;

Qy          1 MPGVCDRAPDFLSPSEDQVLRPALGSSVALNCTAWVVSGPHCSLPSVQWLKDGLPLGIGG 60
              | |||| ||:||||||| |  |||  |||||||||| | | |  |||||||||| || |
Db          1 MAGVCDMAPNFLSPSEDQALGLALGREVALNCTAWVFSRPQCPQPSVQWLKDGLALGNGS 60

Qy         61 HYSLHEYSWVKANLSEVLVSSVLGVNVTSTEVYGAFTCSIQNISFSSFTLQRAGPTSHVA 120
              |:||||  || || ||: ||||| :|:|: | || ||||: |:|  |||| ||||  |||
Db         61 HFSLHEDFWVSANFSEI-VSSVLVLNLTNAEDYGTFTCSVWNVSSHSFTLWRAGPAGHVA 119

Qy        121 AVLASLLVLLALLLAALLYVKCRLNVLLWYQDAYGEVEINDGKLYDAYVSYSDCPEDRKF 180
              |||||||||: ||| ||||||||:|||||| | |||:|||||||||||||||||||||||
Db        120 AVLASLLVLVVLLLVALLYVKCRLNMLLWYQDTYREVEMNDGKLYDAYVSYSDCPEDRKF 179

Qy        181 VNFILKPQLERRRGYKLFLDDRDLLPRAEPSADLLVNLSRCRRLIVVLSDAFLSRAWCSH 240
              ||||||||||||||||||:||||||||: |||||||||||||||||||||||||| |||
Db        180 VNFILKPQLERRRGYKLFLEDRDLLPRADASADLLVNLSRCRRLIVVLSDAFLSRPWCSQ 239

Qy        241 SFREGLCRL----LELTRRPIFITFEGQRRDPAHPALRL---LRQHRHLVTLLLWRPGSV 293
              ||||           :|:  |:        |      |   |     |  ||||||:||:||||
Db        240 SFREDCAAYWSSPADLSSSPL-------RARGVSPYTLLSGSCRHDRHLVTLVLWKPGSV 292

Qy        294 TPSSDFWKEVQLALPRKVRYRPVEGDPQTQLQDDKDPMLILRGRVPEGRALDSEVDPDPE 353
              |||||||||:||||||||:||||||||||:|||  :|| ::||:|||||
Db        293 TPSSDFWKELQLALPRKVQYRPVEGDPQTRLQDDKDPMLIVRGRAAQGRGMESELDPDPE 352

Qy        354 GDLGVRGPVFGEPSAPPHTSGVSLGESRSSEVDVSDLGSRNYSARTDFYCLVSKDDM 410
              |||||||||||||  |   : : :|||  ||:||||||||||||||||||:||:
Db        353 GDLGVRGPVFGEPPTPLQETRICIGESHGSEMDVSDLGSRNYSARTDFYCLVSEDDV 409


RESULT 5
IPL2_HUMAN
ID   IPL2_HUMAN     STANDARD;      PRT;   686 AA.
AC   Q9NP60; Q9NZN0;
DT   10-OCT-2003 (Rel. 42, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   X-linked interleukin-1 receptor accessory protein-like 2 precursor
DE   (IL1RAPL-2 related protein) (Interleukin-1 receptor 9) (IL-1R9) (IL-1
DE   receptor accessory protein-like 2) (Three immunoglobulin domain-
DE   containing IL-1 receptor-related 1) (TIGIRR-1).
GN   Name=IL1RAPL2; Synonyms=IL1R9;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RC   TISSUE=Fetal brain;
RX   MEDLINE=20487552; PubMed=11031108; DOI=10.1006/geno.2000.6328;
RA   Sana T.R., Debets R., Timans J.C., Bazan J.F., Kastelein R.A.;
RT   "Computational identification, cloning, and characterization of IL-
RT   1R9, a novel interleukin-1 receptor-like gene encoded over an
RT   unusually large interval of human chromosome Xq22.2-q22.3.";
RL   Genomics 69:252-262(2000).
RN   [2]
RP   SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RC   TISSUE=Liver;
RX   MEDLINE=20459050; PubMed=10882729; DOI=10.1074/jbc.M004077200;
RA   Born T.L., Smith D.E., Garka K.E., Renshaw B.R., Bertles J.S.,
RA   Sims J.E.;
RT   "Identification and characterization of two members of a novel class
RT   of the interleukin-1 receptor (IL-1R) family. Delineation of a new
RT   class of IL-1R-related proteins based on signaling.";
RL   J. Biol. Chem. 275:29946-29954(2000).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RA   Grabowski M., Lorenz B., Hubel R., Strom T.M.;
```

```
RT   "A gene (IL1RAPL-2) with 61% identity to IL1RAPL maps to Xq22.2.";
RL   Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RC   TISSUE=Brain;
RX   MEDLINE=21472256; PubMed=11587848; DOI=10.1016/S0378-1119(01)00659-X;
RA   Ferrante M.I., Ghiani M., Bulfone A., Franco B.;
RT   "IL1RAPL2 maps to Xq22 and is specifically expressed in the central
RT   nervous system.";
RL   Gene 275:217-221(2001).
RN   [5]
RP   SEQUENCE OF 29-686 FROM N.A.
RX   MEDLINE=20218565; PubMed=10757639; DOI=10.1038/sj.ejhg.5200415;
RA   Jin H., Viswesvaraiah R., Gardner R.J., Roberts R.G.;
RT   "Two novel members of the interleukin-1 receptor gene family, one
RT   deleted in Xp22.3-Xp21.3 mental retardation.";
RL   Eur. J. Hum. Genet. 8:87-94(2000).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- TISSUE SPECIFICITY: Detected at low levels in fetal and adult
CC       brain, in particular in the frontal lobe, temporal lobe and
CC       cerebellum. Detected at very low levels in skin, liver, fetal
CC       ovary and in placenta.
CC   -!- SIMILARITY: Belongs to the interleukin-1 receptor family.
CC   -!- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC   -!- SIMILARITY: Contains 1 TIR domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF212016; AAF61307.1; -.
DR   EMBL; AF284436; AAG21370.1; -.
DR   EMBL; AJ272208; CAB86868.1; -.
DR   EMBL; AJ290436; CAB89867.1; -.
DR   EMBL; AF181285; AAF59412.1; -.
DR   HSSP; Q63345; 1PKO.
DR   Genew; HGNC:5997; IL1RAPL2.
DR   MIM; 300277; -.
DR   GO; GO:0016020; C:membrane; ISS.
DR   GO; GO:0004908; F:interleukin-1 receptor activity; TAS.
DR   GO; GO:0007417; P:central nervous system development; TAS.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR004075; IL1_receptor1.
DR   InterPro; IPR004077; IL1_receptorIIp.
DR   InterPro; IPR000157; TIR.
DR   Pfam; PF00047; ig; 3.
DR   Pfam; PF01582; TIR; 1.
DR   PRINTS; PR01539; INTRLEUKN1R2.
DR   PRINTS; PR01537; INTRLKN1R1F.
DR   SMART; SM00409; IG; 3.
DR   SMART; SM00255; TIR; 1.
DR   PROSITE; PS50835; IG_LIKE; 3.
DR   PROSITE; PS50104; TIR; 1.
KW   Glycoprotein; Immunoglobulin domain; Receptor; Repeat; Signal;
KW   Transmembrane.
FT   SIGNAL        1     16       Potential.
FT   CHAIN        17    686       X-linked interleukin-1 receptor accessory
FT                                 protein-like 2.
FT   DOMAIN       17    354       Extracellular (Potential).
FT   TRANSMEM    355    375       Potential.
FT   DOMAIN      376    686       Cytoplasmic (Potential).
FT   DOMAIN       18    132       Ig-like C2-type 1.
FT   DOMAIN      141    232       Ig-like C2-type 2.
FT   DOMAIN      239    347       Ig-like C2-type 3.
FT   DOMAIN      400    559       TIR.
FT   DISULFID     53    116       Potential.
FT   DISULFID    162    214       Potential.
FT   DISULFID    265    331       Potential.
FT   CARBOHYD     63     63       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    120    120       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    136    136       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    211    211       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    328    328       N-linked (GlcNAc. . .) (Potential).
SQ   SEQUENCE   686 AA;  78669 MW;  E400F7ECD186957C CRC64;

  Query Match             12.5%;  Score 267.5;  DB 1;  Length 686;
  Best Local Similarity   26.7%;  Pred. No. 1e-13;
  Matches  111;  Conservative   75;  Mismatches  152;  Indels   77;  Gaps   21;

Qy          6 DRAPDFLSPSEDQ--VLRPALGSSVALNCTAWV-VSGPHCSLPSVQWLK-DGLPLGIGGH 61
              |: |  | | |:|  |:   ||  : : | |:   ||   | | : |:| :     : ||
Db        237 DKPPKPLFPMENQPSVIDVQLGKPLNIPCKAFFGFSGE--SGPMIYWMKGEKFIEELAGH 294

Qy         62 YSLHEYSWVKANLSE------VLVSSVLGVNVTSTEVYGAFTCSIQNIS---FSSFTLQR 112
                  |   :| :| |       ::  ||:  :: :      :|| ::| :    :|  |::
Db        295 IREGEIRLLKEHLGEKEVELALIFDSVVEADLAN------YTCHVENRNGRKHASVLLRK 348

Qy        113 AGPTSHV--AAVLASLLVLLALLLAALLYVKCRLNVLLWYQDAYGEVEIN-DGKLYDAYV 169
                :  | | :: :|| ||:  ::|    : ::|:|:  :|  | | | | ||||:
Db        349 KDLIYKIELAGGLGAIFLLLVLLV--VIYKCYNIELMLFYRQHFGADETNDDNKEYDAYL 406

Qy        170 SYS-------DC--PEDRKFVNFILKPQLERRRGYKLFLDDRDLLPRAEPSADLLVNLSR 220
              ||:       ||  ||: :|   :|   ||:  |||||: :|||:|      ||   : :
Db        407 SYTKVDQDTLDCDNPEEEQFALEVLPDVLEKHYGYKLFIPERDLIPSGTYMEDLTRYVEQ 466

Qy        221 CRRLIVVLS-DAFLSRAWCSHSFREGLCRLL---ELTRRPIFITFEGQRRDPAHPALRLL 276
               ||||:||: |  | | |          |  :|   |:  : | |     :       :  |
Db        467 SRRLIIVLTPDYILRRGWSIFELESRLHNMLVSGEI--KVILIECTELKGKVNCQEVESL 524

Qy        277 RQHRHLVTLLLWRPGSVTP--SSDFWKEVQLALPRKVRYRPVEGDPQTQLQDDKDPMLIL 334
              ::   |::|: |: || :   :| ||| :   :| |                 |  ||
Db        525 KRSIKLLSLIKWK-GSKSSKLNSKFWKHLVYEMPIK----------------KKEML-- 564

Qy        335 RGRVPEGRALDSEVDPDPEGDLGVRGPVFGEPSAPPHTSGVSLGESRSSEVDVSD 389
                  |    |||        : |: | :   ||    ::  :|    ||: |: :
Db        565 ----PRCHVLDS-------AEQGLFGELQPIPSIAMTSTSATL---VSSQADLPE 605


RESULT 6
IPL2_MOUSE
ID   IPL2_MOUSE     STANDARD;      PRT;   686 AA.
AC   Q9ERS6; Q9ER66;
DT   10-OCT-2003 (Rel. 42, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   X-linked interleukin-1 receptor accessory protein-like 2 precursor
DE   (IL1RAPL-2 related protein) (TIGIRR-1).
GN   Name=Il1rapl2;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   TISSUE=Brain, and Liver;
RX   MEDLINE=20459050; PubMed=10882729; DOI=10.1074/jbc.M004077200;
RA   Born T.L., Smith D.E., Garka K.E., Renshaw B.R., Bertles J.S.,
RA   Sims J.E.;
RT   "Identification and characterization of two members of a novel class
RT   of the interleukin-1 receptor (IL-1R) family. Delineation of a new
RT   class of IL-1R-related proteins based on signaling.";
RL   J. Biol. Chem. 275:29946-29954(2000).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM 2), AND TISSUE SPECIFICITY.
RC   TISSUE=Brain;
RX   MEDLINE=21472256; PubMed=11587848; DOI=10.1016/S0378-1119(01)00659-X;
RA   Ferrante M.I., Ghiani M., Bulfone A., Franco B.;
RT   "IL1RAPL2 maps to Xq22 and is specifically expressed in the central
RT   nervous system.";
RL   Gene 275:217-221(2001).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
```

```
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=1;
CC         IsoId=Q9ERS6-1; Sequence=Displayed;
CC       Name=2;
CC         IsoId=Q9ERS6-2; Sequence=VSP_008056, VSP_008057;
CC         Note=may be due to an intron retention;
CC   -!- TISSUE SPECIFICITY: Detected in fetal brain after day 12.5, in
CC       particular in parts of the diencephalon and in the basal plate of
CC       the spinal cord. In postnatal brain detected in cerebral cortex,
CC       olfactory bulb, in the CA1 region of the hippocampus and in
CC       Purkinje cells of the Xth cerebellar lobule.
CC   -!- SIMILARITY: Belongs to the interleukin-1 receptor family.
CC   -!- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC   -!- SIMILARITY: Contains 1 TIR domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF284437; AAG21371.1; -.
DR   EMBL; AJ277831; CAC10559.1; -.
DR   HSSP; Q63345; 1PKO.
DR   MGD; MGI:1913106; Il1rapl2.
DR   GO; GO:0016020; C:membrane; ISS.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR004075; IL1_receptor1.
DR   InterPro; IPR004077; IL1_receptorIIp.
DR   InterPro; IPR000157; TIR.
DR   Pfam; PF00047; ig; 3.
DR   Pfam; PF01582; TIR; 1.
DR   PRINTS; PR01539; INTRLEUKN1R2.
DR   PRINTS; PR01537; INTRLKN1R1F.
DR   SMART; SM00409; IG; 3.
DR   SMART; SM00255; TIR; 1.
DR   PROSITE; PS50835; IG_LIKE; 3.
DR   PROSITE; PS50104; TIR; 1.
KW   Alternative splicing; Glycoprotein; Immunoglobulin domain; Receptor;
KW   Repeat; Signal; Transmembrane.
FT   SIGNAL        1     16       Potential.
FT   CHAIN        17    686       X-linked interleukin-1 receptor accessory
FT                                 protein-like 2.
FT   DOMAIN       17    354       Extracellular (Potential).
FT   TRANSMEM     355   375       Potential.
FT   DOMAIN       376   686       Cytoplasmic (Potential).
FT   DOMAIN        32   132       Ig-like C2-type 1.
FT   DOMAIN       141   232       Ig-like C2-type 2.
FT   DOMAIN       239   347       Ig-like C2-type 3.
FT   DOMAIN       400   559       TIR.
FT   DISULFID      53   116       Potential.
FT   DISULFID     162   214       Potential.
FT   DISULFID     265   331       Potential.
FT   CARBOHYD      63    63       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD     120   120       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD     136   136       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD     211   211       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD     328   328       N-linked (GlcNAc. . .) (Potential).
FT   VARSPLIC     350   354       DLIYK -> GILFS (in isoform 2).
FT                                 /FTId=VSP_008056.
FT   VARSPLIC     355   686       Missing (In isoform 2).
FT                                 /FTId=VSP_008057.
SQ   SEQUENCE   686 AA;  78797 MW;  36160D1CDE9B8264 CRC64;

  Query Match             12.2%;  Score 261;  DB 1;  Length 686;
  Best Local Similarity   28.9%;  Pred. No. 3.7e-13;
  Matches   98;  Conservative   65;  Mismatches  126;  Indels   50;  Gaps   18;

Qy          6 DRAPDFLSPSEDQ--VLRPALGSSVALNCTAWV-VSGPHCSLPSVQWLK-DGLPLGIGGH 61
              |: |  | | |:|  |:   ||  : : | |:   ||   | | : |:| :     : ||
Db        237 DKPPKPLFPMENQPSVIDVQLGKPLNIPCKAFFGFSGE--SGPMIYWMKGEKFIEELAGH 294

Qy         62 YSLHEYSWVKANLSE------VLVSSVLGVNVTSTEVYGAFTCSIQNIS---FSSFTLQR 112
                 |  :| :| |       ::  ||:  :: :      :|| ::| :    :|  |::
Db        295 IREGEIRLLKEHLGEKEVELTLIFDSVVEADLAN------YTCHVENRNGRKHASVLLRK 348

Qy        113 AGPTSHV--AAVLASLLVLLALLLAALLYVKCRLNVLLWYQDAYGEVE-INDGKLYDAYV 169
                :  |  | :: :|| |||  ::|    : ::|:|:  :|  |  :| | ||||:
Db        349 KDLIYKIELAGGLGAIFLLLILLL--VVYKCYNIELMLFYRQRFGGDETTDDNKEYDAYL 406

Qy        170 SYS-------DC--PEDRKFVNFILKPQLERRRGYKLFLDDRDLLPRAEPSADLLVNLSR 220
              ||:       ||   |: :|   ||   ||:  |||||: :|||:    ||   : :|:|
Db        407 SYTKVDQDTLDCDNTEEEQFALEILPDVLEKHYGYKLFIPERDLI----PSGTYIEDLTR 462

Qy        221 C----RRLIVVLS-DAFLSRAWCSHSFREGLCRLL---ELTRRPIFITFEGQRRDPAHPA 272
              |    ||||:||: |  | | |          |  :|   |:  : | |     :
Db        463 CVEQSRRLIIVLTPDYILRRGWSIFELESRLHNMLVSGEI--KVILIECTELKGKVNCQE 520

Qy        273 LRLLRQHRHLVTLLLWR-PGSVTPSSDFWKEVQLALPRK 310
              :  |: :  |::|: |: | |   :| ||| :   :| |
Db        521 VESLKHNIKLLSLIKWKGPKSSKLNSKFWKHLVYEMPIK 559
```

```
RESULT 7
IPL1_MOUSE
ID   IPL1_MOUSE     STANDARD;      PRT;   695 AA.
AC   P59823;
DT   10-OCT-2003 (Rel. 42, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   X-linked interleukin-1 receptor accessory protein-like 1 precursor
DE   (IL1RAPL-1).
GN   Name=Il1rapl1;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   CONCEPTUAL TRANSLATION.
RA   Hinz U.;
RL   Unpublished observations (AUG-2003).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- SIMILARITY: Belongs to the interleukin-1 receptor family.
CC   -!- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC   -!- SIMILARITY: Contains 1 TIR domain.
CC   -!- CAUTION: This is a conceptual translation using ESTs and genomic
CC       sequences, in analogy to ortholog sequences.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AL844900; -; NOT_ANNOTATED_CDS.
DR   EMBL; AL672059; -; NOT_ANNOTATED_CDS.
DR   EMBL; AK081272; -; NOT_ANNOTATED_CDS.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003598; Ig_c2.
DR   InterPro; IPR004075; IL1_receptor1.
DR   InterPro; IPR000157; TIR.
DR   Pfam; PF00047; ig; 2.
DR   Pfam; PF01582; TIR; 1.
DR   PRINTS; PR01537; INTRLKN1R1F.
DR   SMART; SM00408; IGc2; 1.
DR   SMART; SM00255; TIR; 1.
DR   PROSITE; PS50835; IG_LIKE; 3.
DR   PROSITE; PS50104; TIR; 1.
KW   Glycoprotein; Immunoglobulin domain; Receptor; Repeat; Signal;
KW   Transmembrane.
```

```
FT   SIGNAL        1     24       Potential.
FT   CHAIN        25    695       X-linked interleukin-1 receptor accessory
FT                                 protein-like 1.
FT   DOMAIN       25    357       Extracellular (Potential).
FT   TRANSMEM    358    378       Potential.
FT   DOMAIN      379    695       Cytoplasmic (Potential).
FT   DOMAIN       25    134       Ig-like C2-type 1.
FT   DOMAIN      143    232       Ig-like C2-type 2.
FT   DOMAIN      242    350       Ig-like C2-type 3.
FT   DOMAIN      403    561       TIR.
FT   DISULFID     53    118       Potential.
FT   DISULFID    164    216       Potential.
FT   DISULFID    267    334       Potential.
FT   CARBOHYD     63     63       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    122    122       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    138    138       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    213    213       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    264    264       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    331    331       N-linked (GlcNAc. . .) (Potential).
SQ   SEQUENCE   695 AA;  79630 MW;  5FE34F204E5908B7 CRC64;

  Query Match           12.0%;  Score 257;  DB 1;  Length 695;
  Best Local Similarity   26.2%;  Pred. No. 8.2e-13;
  Matches   95;  Conservative   64;  Mismatches  145;  Indels   58;  Gaps   17;

Qy          6 DRAPDFLSPSEDQ--VLRPALGSSVALNCTAWVVSGPHCSLPSVQWLKDGLPLGIGGHYS 63
              |: |  | | | :  |    || |  | | |:       | | : |:|     |
Db        239 DKPPKLLYPMESKLTVQETQLGGSANLTCRAFFGYSGDVS-PLIYWMK--------GEKF 289

Qy         64 LHEYS----W------VKANLSEVLVSSVLGVNVTSTEVYGAFTCSIQNIS---FSSFTL 110
              : :      |      :| :| |  ||  | |:       | ::| ::| :    :|  |
Db        290 IEDLDENRVWESDIRILKEHLGEQEVSISLIVDSVEEGDLGNYSCYVENGNGRRHASVLL 349

Qy        111 QRAGPTSHVAAVLASLLVLLALLLAALLYVKC-RLNVLLWYQDAYGEVEIN-DGKLYDAY 168
               :    :   :   |  :| ||: ::   || :: ::|:|:: :|  |:: | | ||||
Db        350 HKR-ELMYTVELAGGLGAILLLLICSVTIYKCYKIEIMLFYRNHFGAEELDGDNKDYDAY 408

Qy        169 VSYS---------DCPEDRKFVNFILKPQLERRRGYKLFLDDRDLLPRAEPSADLLVNLS 219
              :||:         :  |: :|   ||   ||:  ||||: ||||:|       : :::
Db        409 LSYTKVDPDQWNQETGEEERFALEILPDMLEKHYGYKLFIPDRDLIPTGN-----IEDVA 463

Qy        220 RC----RRLIVVLSDAF-LSRAWCSHSFREGLCRLLELTR-RPIFITFEGQRRDPAHPAL 273
              ||    :|||:|::  : : | |         |  :|     : | |     |    :  :
Db        464 RCVDQSKRLIIVMTPNYVVRRGWSIFELETRLRNMLVTGEIKVILIECSELRGIMNYQEV 523

Qy        274 RLLRQHRHLVTLLLWR-PGSVTPSSDFWKEVQLALPRKVRYRPV---------EGDPQTQ 323
                |:    |:|:: |  |     :| ||| :|  :| | |  |:         |  |  :
Db        524 EALKHTIKLLTVIKWHGPKCNKLNSKFWKRLQYEMPFK-RIEPITHEQALDVSEQGPFGE 582

Qy        324 LQ 325
              ||
Db        583 LQ 584


RESULT 8
IPL1_HUMAN
ID   IPL1_HUMAN     STANDARD;      PRT;   696 AA.
AC   Q9NZN1; Q9UJ53;
DT   10-OCT-2003 (Rel. 42, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   X-linked interleukin-1 receptor accessory protein-like 1 precursor
DE   (IL1RAPL-1) (Oligophrenin-4) (Three immunoglobulin domain-containing
DE   IL-1 receptor-related 2) (TIGIRR-2).
GN   Name=IL1RAPL1; Synonyms=OPHN4;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., DISEASE, TISSUE SPECIFICITY, AND ALTERNATIVE
RP   SPLICING.
RC   TISSUE=Fetal brain;
RX   MEDLINE=99400548; PubMed=10471494; DOI=10.1038/12623;
RA   Carrie A., Jun L., Bienvenu T., Vinet M.-C., McDonell N., Couvert P.,
RA   Zemni R., Cardona A., Van Buggenhout G., Frints S., Hamel B.,
RA   Moraine C., Ropers H.-H., Strom T., Howell G.R., Whittaker A.,
RA   Ross M.T., Kahn A., Fryns J.-P., Beldjord C., Marynen P., Chelly J.;
RT   "A new member of the IL-1 receptor family highly expressed in
RT   hippocampus and involved in X-linked mental retardation.";
RL   Nat. Genet. 23:25-31(1999).
RN   [2]
RP   SEQUENCE FROM N.A., AND DISEASE.
RX   MEDLINE=20218565; PubMed=10757639; DOI=10.1038/sj.ejhg.5200415;
RA   Jin H., Gardner R.J., Viswesvaraiah R., Muntoni F., Roberts R.G.;
RT   "Two novel members of the interleukin-1 receptor gene family, one
RT   deleted in Xp22.3-Xp21.3 mental retardation.";
RL   Eur. J. Hum. Genet. 8:87-94(2000).
RN   [3]
RP   SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RC   TISSUE=Brain, and Testis;
RX   MEDLINE=20459050; PubMed=10882729; DOI=10.1074/jbc.M004077200;
RA   Born T.L., Smith D.E., Garka K.E., Renshaw B.R., Bertles J.S.,
RA   Sims J.E.;
RT   "Identification and characterization of two members of a novel class
RT   of the interleukin-1 receptor (IL-1R) family. Delineation of a new
RT   class of IL-1R-related proteins based on signaling.";
RL   J. Biol. Chem. 275:29946-29954(2000).
RN   [4]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22763540; PubMed=12777533; DOI=10.1093/molbev/msg134;
RA   Kitano T., Schwarz C., Nickel B., Paabo S.;
RT   "Gene diversity patterns at 10 X-chromosomal loci in humans and
RT   chimpanzees.";
RL   Mol. Biol. Evol. 20:1281-1289(2003).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=1;
CC         Comment=A number of isoforms are produced;
CC       Name=1;
CC         IsoId=Q9NZN1-1; Sequence=Displayed;
CC   -!- TISSUE SPECIFICITY: Detected at low levels in heart, skeletal
CC       muscle, ovary, skin, and in amygdala, caudate nucleus, corpus
CC       callosum, hippocampus, substantia nigra and thalamus. Detected at
CC       very low levels in tonsil, prostate, testis, small intestine,
CC       placenta, colon and fetal liver.
CC   -!- DISEASE: Defects in IL1RAPL1 are a cause of nonspecific X-linked
CC       mental retardation (MRX34) [MIM:300426]. Several forms of X-linked
CC       mental retardation have been described; they are classified as
CC       non-specific if there are no other characteristic clinical
CC       symptoms.
CC   -!- SIMILARITY: Belongs to the interleukin-1 receptor family.
CC   -!- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC   -!- SIMILARITY: Contains 1 TIR domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AJ243874; CAB56046.1; -.
DR   EMBL; AF181284; AAF59411.1; -.
DR   EMBL; AF284435; AAG21369.1; -.
DR   EMBL; AB102650; BAC81119.1; -.
DR   HSSP; P11362; 1EVT.
DR   Genew; HGNC:5996; IL1RAPL1.
DR   MIM; 300206; -.
DR   MIM; 300426; -.
DR   GO; GO:0016020; C:membrane; ISS.
DR   GO; GO:0004872; F:receptor activity; TAS.
DR   GO; GO:0007611; P:learning and/or memory; TAS.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR004075; IL1_receptor1.
```

```
DR   InterPro; IPR000157; TIR.
DR   Pfam; PF00047; ig; 2.
DR   Pfam; PF01582; TIR; 1.
DR   PRINTS; PR01537; INTRLKN1R1F.
DR   PROSITE; PS50835; IG_LIKE; 3.
DR   PROSITE; PS50104; TIR; 1.
KW   Alternative splicing; Glycoprotein; Immunoglobulin domain; Receptor;
KW   Repeat; Signal; Transmembrane.
FT   SIGNAL       1     18       Potential.
FT   CHAIN       19    696       X-linked interleukin-1 receptor accessory
FT                               protein-like 1.
FT   DOMAIN      19    357       Extracellular (Potential).
FT   TRANSMEM   358    378       Potential.
FT   DOMAIN     379    696       Cytoplasmic (Potential).
FT   DOMAIN      19    134       Ig-like C2-type 1.
FT   DOMAIN     143    232       Ig-like C2-type 2.
FT   DOMAIN     242    350       Ig-like C2-type 3.
FT   DOMAIN     403    562       TIR.
FT   DISULFID    53    118       Potential.
FT   DISULFID   164    216       Potential.
FT   DISULFID   267    334       Potential.
FT   CARBOHYD    63     63       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   122    122       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   138    138       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   213    213       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   264    264       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   331    331       N-linked (GlcNAc. . .) (Potential).
SQ   SEQUENCE   696 AA;  79968 MW;  9B7A0B503D73CCA9 CRC64;

  Query Match             11.9%;  Score 256.5;  DB 1;  Length 696;
  Best Local Similarity   26.5%;  Pred. No. 9.1e-13;
  Matches   95;  Conservative   64;  Mismatches  149;  Indels   51;  Gaps   15;

Qy          6 DRAPDFLSPSEDQ--VLRPALGSSVALNCTAWVVSGPHCSLPSVQWLKDGLPLGIGGHYS 63
              |: |  | | | :  :    || |  | | |:       | | : |:|       |
Db        239 DKPPKLLYPMESKLTIQETQLGDSANLTCRAFFGYSGDVS-PLIYWMK--------GEKF 289

Qy         64 LHEYS----W------VKANLSEVLVSSVLGVNVTSTEVYGAFTCSIQNIS---FSSFTL 110
              : :      |      :| :| |  ||  | |:       | ::| ::| :    :|  |
Db        290 IEDLDENRVWESDIRILKEHLGEQEVSISLIVDSVEEGDLGNYSCYVENGNGRRHASVLL 349

Qy        111 QRAG--PTSHVAAVLASLLVLLALLLAALLYVKCRLNVLLWYQDAYGEVEIN-DGKLYDA 167
               :     |  :|  | ::|:||  |:   :|   :: ::|:|:: :|  |:: | | |||
Db        350 HKRELMYTVELAGGLGAILLLLVCLVT--IYKCYKIEIMLFYRNHFGAEELDGDNKDYDA 407

Qy        168 YVSYS---------DCPEDRKFVNFILKPQLERRRGYKLFLDDRDLLPRAEPSADLLVNL 218
              |:||:         :  |: :|   ||   ||:  |||||: ||||:|      |:   :
Db        408 YLSYTKVDPDQWNQETGEEERFALEILPDMLEKHYGYKLFIPDRDLIPTGTYIEDVARCV 467

Qy        219 SRCRRLIVVLSDAF-LSRAWCSHSFREGLCRLLELTR-RPIFITFEGQRRDPAHPALRLL 276
               : :|||:|::  : : | |        |  :|     : | |     |    :  :  |
Db        468 DQSKRLIIVMTPNYVVRRGWSIFELETRLRNMLVTGEIKVILIECSELRGIMNYQEVEAL 527

Qy        277 RQHRHLVTLLLWR-PGSVTPSSDFWKEVQLALPRKVRYRPV---------EGDPQTQLQ 325
               :    |:|:: |  |      :| ||| :|  :| | |  |:         |  |  :||
Db        528 KHTIKLLTVIKWHGPKCNKLNSKFWKRLQYEMPFK-RIEPITHEQALDVSEQGPFGELQ 585


RESULT 9
IPL1_PANTR
ID   IPL1_PANTR     STANDARD;      PRT;   696 AA.
AC   P60029;
DT   29-MAR-2004 (Rel. 43, Created)
DT   29-MAR-2004 (Rel. 43, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   X-linked interleukin-1 receptor accessory protein-like 1 precursor
DE   (IL1RAPL-1) (Oligophrenin-4).
GN   Name=IL1RAPL1; Synonyms=OPHN4;
OS   Pan troglodytes (Chimpanzee).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX   NCBI_TaxID=9598;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22763540; PubMed=12777533; DOI=10.1093/molbev/msg134;
RA   Kitano T., Schwarz C., Nickel B., Paabo S.;
RT   "Gene diversity patterns at 10 X-chromosomal loci in humans and
RT   chimpanzees.";
RL   Mol. Biol. Evol. 20:1281-1289(2003).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- SIMILARITY: Belongs to the interleukin-1 receptor family.
CC   -!- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC   -!- SIMILARITY: Contains 1 TIR domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AB102651; BAC81120.1; -.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR000157; TIR.
DR   Pfam; PF00047; ig; 2.
DR   Pfam; PF01582; TIR; 1.
DR   PROSITE; PS50835; IG_LIKE; 3.
DR   PROSITE; PS50104; TIR; 1.
KW   Glycoprotein; Immunoglobulin domain; Receptor; Repeat; Signal;
KW   Transmembrane.
FT   SIGNAL       1     18       Potential.
FT   CHAIN       19    696       X-linked interleukin-1 receptor accessory
FT                               protein-like 1.
FT   DOMAIN      19    357       Extracellular (Potential).
FT   TRANSMEM   358    378       Potential.
FT   DOMAIN     379    696       Cytoplasmic (Potential).
FT   DOMAIN      19    134       Ig-like C2-type 1.
FT   DOMAIN     143    232       Ig-like C2-type 2.
FT   DOMAIN     242    350       Ig-like C2-type 3.
FT   DOMAIN     403    562       TIR.
FT   DISULFID    53    118       Potential.
FT   DISULFID   164    216       Potential.
FT   DISULFID   267    334       Potential.
FT   CARBOHYD    63     63       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   122    122       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   138    138       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   213    213       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   264    264       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   331    331       N-linked (GlcNAc. . .) (Potential).
SQ   SEQUENCE   696 AA;  79968 MW;  9B7A0B503D73CCA9 CRC64;

  Query Match             11.9%;  Score 256.5;  DB 1;  Length 696;
  Best Local Similarity   26.5%;  Pred. No. 9.1e-13;
  Matches   95;  Conservative   64;  Mismatches  149;  Indels   51;  Gaps   15;

Qy          6 DRAPDFLSPSEDQ--VLRPALGSSVALNCTAWVVSGPHCSLPSVQWLKDGLPLGIGGHYS 63
              |: |  | | | :  :    || |  | | |:       | | : |:|       |
Db        239 DKPPKLLYPMESKLTIQETQLGDSANLTCRAFFGYSGDVS-PLIYWMK--------GEKF 289

Qy         64 LHEYS----W------VKANLSEVLVSSVLGVNVTSTEVYGAFTCSIQNIS---FSSFTL 110
              : :      |      :| :| |  ||  | |:       | ::| ::| :    :|  |
Db        290 IEDLDENRVWESDIRILKEHLGEQEVSISLIVDSVEEGDLGNYSCYVENGNGRRHASVLL 349

Qy        111 QRAG--PTSHVAAVLASLLVLLALLLAALLYVKCRLNVLLWYQDAYGEVEIN-DGKLYDA 167
               :     |  :|  | ::|:||  |:   :|   :: ::|:|:: :|  |:: | | |||
Db        350 HKRELMYTVELAGGLGAILLLLVCLVT--IYKCYKIEIMLFYRNHFGAEELDGDNKDYDA 407

Qy        168 YVSYS---------DCPEDRKFVNFILKPQLERRRGYKLFLDDRDLLPRAEPSADLLVNL 218
              |:||:         :  |: :|   ||   ||:  |||||: ||||:|      |:   :
Db        408 YLSYTKVDPDQWNQETGEEERFALEILPDMLEKHYGYKLFIPDRDLIPTGTYIEDVARCV 467

Qy        219 SRCRRLIVVLSDAF-LSRAWCSHSFREGLCRLLELTR-RPIFITFEGQRRDPAHPALRLL 276
               : :|||:|::  : : | |        |  :|     : | |     |    :  :  |
Db        468 DQSKRLIIVMTPNYVVRRGWSIFELETRLRNMLVTGEIKVILIECSELRGIMNYQEVEAL 527
```

```
Qy        277 RQHRHLVTLLLWR-PGSVTPSSDFWKEVQLALPRKVRYRPV---------EGDPQTQLQ 325
              :    |:|:: |  |     :| ||| :|  :| | |  |:         |  |  :||
Db        528 KHTIKLLTVIKWHGPKCNKLNSKFWKRLQYEMPFK-RIEPITHEQALDVSEQGPFGELQ 585


RESULT 10
IPL1_RAT
ID   IPL1_RAT       STANDARD;      PRT;   696 AA.
AC   P59824;
DT   10-OCT-2003 (Rel. 42, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   X-linked interleukin-1 receptor accessory protein-like 1 precursor
DE   (IL1RAPL-1).
GN   Name=Il1rapl1;
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Wistar; TISSUE=Hippocampus;
RA   Boda B., Parisi L., Muller D.;
RT   "Full length cloning of rat IL1RAPL in the hippocampus.";
RL   Submitted (JAN-2003) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- SIMILARITY: Belongs to the interleukin-1 receptor family.
CC   -!- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC   -!- SIMILARITY: Contains 1 TIR domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AY216593; AAO62634.1; -.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003598; Ig_c2.
DR   InterPro; IPR004075; IL1_receptor1.
DR   InterPro; IPR000157; TIR.
DR   Pfam; PF00047; ig; 2.
DR   Pfam; PF01582; TIR; 1.
DR   PRINTS; PR01537; INTRLKN1R1F.
DR   SMART; SM00408; IGc2; 1.
DR   SMART; SM00255; TIR; 1.
DR   PROSITE; PS50835; IG_LIKE; 3.
DR   PROSITE; PS50104; TIR; 1.
KW   Glycoprotein; Immunoglobulin domain; Receptor; Repeat; Signal;
KW   Transmembrane.
FT   SIGNAL        1     24       Potential.
FT   CHAIN        25    696       X-linked interleukin-1 receptor accessory
FT                                 protein-like 1.
FT   DOMAIN       25    357       Extracellular (Potential).
FT   TRANSMEM     358   378       Potential.
FT   DOMAIN       379   696       Cytoplasmic (Potential).
FT   DOMAIN       25    134       Ig-like C2-type 1.
FT   DOMAIN       143   232       Ig-like C2-type 2.
FT   DOMAIN       242   350       Ig-like C2-type 3.
FT   DOMAIN       403   562       TIR.
FT   DISULFID     53    118       Potential.
FT   DISULFID     164   216       Potential.
FT   DISULFID     267   334       Potential.
FT   CARBOHYD     63    63        N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD     122   122       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD     138   138       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD     213   213       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD     264   264       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD     331   331       N-linked (GlcNAc. . .) (Potential).
SQ   SEQUENCE   696 AA;  79771 MW;  3684FCCD0A0B3FD9 CRC64;
```

```
  Query Match           11.9%;  Score 256.5;  DB 1;  Length 696;
  Best Local Similarity  26.0%;  Pred. No. 9.1e-13;
  Matches   93;  Conservative   64;  Mismatches  152;  Indels   49;  Gaps   15;

Qy          6 DRAPDFLSPSEDQ--VLRPALGSSVALNCTAWVVSGPHCSLPSVQWLKDGLPLGIGGHYS 63
              |: |  | | | :  :    || |  | | |:       | | : |:|       |
Db        239 DKPPKLLYPMESKLTIQETQLGGSANLTCRAFFGYSGDVS-PLIYWMK--------GEKF 289

Qy         64 LHEYS----W------VKANLSEVLVSSVLGVNVTSTEVYGAFTCSIQNIS---FSSFTL 110
              : :      |      :| :| |  ||  | |:       | ::| ::| :    :|  |
Db        290 IEDLDENRVWESDIRILKEHLGEQEVSISLIVDSVEEGDLGNYSCYVENGNGRRHASVLL 349

Qy        111 QRAGPTSHVAAVLASLLVLLALLLAALLYVKC-RLNVLLWYQDAYGEVEIN-DGKLYDAY 168
              :    :   :   | :| ||: ::   || :: ::|:|:: :|  |:: | | ||||
Db        350 HKR-ELMYTVELAGGLGAILLLLVCSVTIYKCYKIEIMLFYRNHFGAEELDGDNKDYDAY 408

Qy        169 VSYS---------DCPEDRKFVNFILKPQLERRRGYKLFLDDRDLLPRAEPSADLLVNLS 219
              :||:         :  |: :|   ||   ||:  |||||: ||||:|      |:   :
Db        409 LSYTKVDPDQWNQETGEEERFALEILPDMLEKHYGYKLFIPDRDLIPTGTYIEDVARCVD 468

Qy        220 RCRRLIVVLSDAF-LSRAWCSHSFREGLCRLLELTR-RPIFITFEGQRRDPAHPALRLLR 277
              : :|||:|::  : : | |       |  :|    : | |      |    :   :  |:
Db        469 QSKRLIIVMTPNYVVRRGWSIFELETRLRNMLVTGEIKVILIECSELRGIMNYQEVEALK 528

Qy        278 QHRHLVTLLLWR-PGSVTPSSDFWKEVQLALPRKVRYRPV---------EGDPQTQLQ 325
                |:|:: |  |     :| ||| :|  :| | |  |:         |  |  :||
Db        529 HTIKLLTVIKWHGPKCNKLNSKFWKRLQYEMPFK-RIEPITHEQALDVSEQGPFGELQ 585


RESULT 11
IPL1_PONPY
ID   IPL1_PONPY     STANDARD;      PRT;   696 AA.
AC   Q7YQL9;
DT   05-JUL-2004 (Rel. 44, Created)
DT   05-JUL-2004 (Rel. 44, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   X-linked interleukin-1 receptor accessory protein-like 1 precursor
DE   (IL1RAPL-1) (Oligophrenin-4).
GN   Name=IL1RAPL1; Synonyms=OPHN4;
OS   Pongo pygmaeus (Orangutan).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.
OX   NCBI_TaxID=9600;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22763540; PubMed=12777533; DOI=10.1093/molbev/msg134;
RA   Kitano T., Schwarz C., Nickel B., Paabo S.;
RT   "Gene diversity patterns at 10 X-chromosomal loci in humans and
RT   chimpanzees.";
RL   Mol. Biol. Evol. 20:1281-1289(2003).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- SIMILARITY: Belongs to the interleukin-1 receptor family.
CC   -!- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC   -!- SIMILARITY: Contains 1 TIR domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AB102652; BAC81121.1; -.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR003598; Ig_c2.
DR   InterPro; IPR000157; TIR.
DR   Pfam; PF00047; ig; 2.
DR   Pfam; PF01582; TIR; 1.
DR   SMART; SM00408; IGc2; 1.
DR   SMART; SM00255; TIR; 1.
DR   PROSITE; PS50835; IG_LIKE; 3.
DR   PROSITE; PS50104; TIR; 1.
```

```
KW   Glycoprotein; Immunoglobulin domain; Receptor; Repeat; Signal;
KW   Transmembrane.
FT   SIGNAL        1     18       Potential.
FT   CHAIN        19    696       X-linked interleukin-1 receptor accessory
FT                                 protein-like 1.
FT   DOMAIN       19    357       Extracellular (Potential).
FT   TRANSMEM    358    378       Potential.
FT   DOMAIN      379    696       Cytoplasmic (Potential).
FT   DOMAIN       19    134       Ig-like 1.
FT   DOMAIN      143    232       Ig-like 2.
FT   DOMAIN      242    350       Ig-like 3.
FT   DOMAIN      403    562       TIR.
FT   DISULFID     53    118       Potential.
FT   DISULFID    164    216       Potential.
FT   DISULFID    267    334       Potential.
FT   CARBOHYD     63     63       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    122    122       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    138    138       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    213    213       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    264    264       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    331    331       N-linked (GlcNAc. . .) (Potential).
SQ   SEQUENCE   696 AA;  79936 MW;  0332FE0C31C2D4E1 CRC64;

  Query Match             11.9%;  Score 255.5;  DB 1;  Length 696;
  Best Local Similarity   26.5%;  Pred. No. 1.1e-12;
  Matches   95;  Conservative   64;  Mismatches  149;  Indels   51;  Gaps   15;

Qy          6 DRAPDFLSPSEDQ--VLRPALGSSVALNCTAWVVSGPHCSLPSVQWLKDGLPLGIGGHYS 63
              |: |  | | |  :   :    || |  | | |:       | | : |:|       |
Db        239 DKPPKLLYPVESKLTIQETQLGDSANLTCRAFFGYSGDVS-PLIYWMK--------GEKF 289

Qy         64 LHEYS----W------VKANLSEVLVSSVLGVNVTSTEVYGAFTCSIQNIS---FSSFTL 110
              : :      |      :| :| |  ||  | |:       | ::| ::| :    :|  |
Db        290 IEDLDENRVWESDIRILKEHLGEQEVSISLIVDSVEEGDLGNYSCYVENGNGRRHASVLL 349

Qy        111 QRAG--PTSHVAAVLASLLVLLALLLAALLYVKCRLNVLLWYQDAYGEVEIN-DGKLYDA 167
               :     |  :|  | ::|:||  |:   :|   :: ::|:|:: :|  |:: | | |||
Db        350 HKRELMYTVELAGGLGAILLLLVCLVT--IYKCYKIEIMLFYRNHFGAEELDGDNKDYDA 407

Qy        168 YVSYS---------DCPEDRKFVNFILKPQLERRRGYKLFLDDRDLLPRAEPSADLLVNL 218
              |:||:         :  |: :|   ||   ||:  |||||: ||||:|       |:   :
Db        408 YLSYTKVDPDQWNQETGEEERFALEILPDMLEKHYGYKLFIPDRDLIPTGTYIEDVARCV 467

Qy        219 SRCRRLIVVLSDAF-LSRAWCSHSFREGLCRLLELTR-RPIFITFEGQRRDPAHPALRLL 276
              : :|||:|::  : : | |         |  :|    : | |      |    :  :  |
Db        468 DQSKRLIIVMTPNYVVRRGWSIFELETRLRNMLVTGEIKVILIECSELRGIMNYQEVEAL 527

Qy        277 RQHRHLVTLLLWR-PGSVTPSSDFWKEVQLALPRKVRYRPV---------EGDPQTQLQ 325
              :   |:|:: |  |     :| ||| :|  :| | |  |:         |  |  :||
Db        528 KHTIKLLTVIKWHGPKCNKLNSKFWKRLQYEMPFK-RIEPITHEQALDVSEQGPFGELQ 585


RESULT 12
IR18_MOUSE
ID   IR18_MOUSE     STANDARD;      PRT;   537 AA.
AC   Q61098;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Interleukin-18 receptor 1 precursor (IL1 receptor-related protein)
DE   (IL-1Rrp).
GN   Name=Il18r1;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A., AND VARIANT ALA-317 DEL.
RX   MEDLINE=96223957; PubMed=8626725; DOI=10.1074/jbc.271.8.3967;
RA   Parnet P., Garka K.E., Bonnert T.P., Dower S.K., Sims J.E.;
RT   "IL-1Rrp is a novel receptor-like molecule similar to the type I
RT   interleukin-1 receptor and its homologues T1/ST2 and IL-1R AcP.";
RL   J. Biol. Chem. 271:3967-3970(1996).
CC   -!- FUNCTION: Receptor for interleukin 18 (IL-18). Binding to the
CC       agonist leads to the activation of NF-kappa-B.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- MISCELLANEOUS: Does not bind interleukin 1 alpha (IL-1A) or
CC       interleukin 1 beta (IL-1B).
CC   -!- SIMILARITY: Belongs to the interleukin-1 receptor family.
CC   -!- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC   -!- SIMILARITY: Contains 1 TIR domain.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; U43673; AAC52437.1; -.
DR   HSSP; Q15399; 1FYV.
DR   MGD; MGI:105383; Il18r1.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR004075; IL1_receptor1.
DR   InterPro; IPR000157; TIR.
DR   Pfam; PF00047; ig; 3.
DR   Pfam; PF01582; TIR; 1.
DR   PRINTS; PR01537; INTRLKN1R1F.
DR   SMART; SM00409; IG; 2.
DR   SMART; SM00255; TIR; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS50104; TIR; 1.
KW   Glycoprotein; Immunoglobulin domain; Polymorphism; Receptor; Repeat;
KW   Signal; Transmembrane.
FT   SIGNAL        1     19       Potential.
FT   CHAIN        20    537       Interleukin-18 receptor 1.
FT   DOMAIN       20    326       Extracellular (Potential).
FT   TRANSMEM    327    347       Potential.
FT   DOMAIN      348    537       Cytoplasmic (Potential).
FT   DOMAIN       33    102       Ig-like C2-type 1.
FT   DOMAIN      134    189       Ig-like C2-type 2.
FT   DOMAIN      212    312       Ig-like C2-type 3.
FT   DOMAIN      370    519       TIR.
FT   DISULFID    141    182       By similarity.
FT   DISULFID    234    296       By similarity.
FT   CARBOHYD     50     50       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD     71     71       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    107    107       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    111    111       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    134    134       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    189    189       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    194    194       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    200    200       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    233    233       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    295    295       N-linked (GlcNAc. . .) (Potential).
FT   VARIANT     317    317       Missing.
SQ   SEQUENCE   537 AA;  61600 MW;  B119PEA403355458 CRC64;

  Query Match             11.8%;  Score 254;  DB 1;  Length 537;
  Best Local Similarity   27.9%;  Pred. No. 1e-12;
  Matches   95;  Conservative   62;  Mismatches  113;  Indels   70;  Gaps   18;

Qy          9 PDFLSPSEDQVLRPALGSSVALNCTA---------WVVSGPHCSLPSVQ--------WLK 51
              |  | |  ::|     ||  | |||:|         | :     | |:||        |:
Db        212 PAILGPKCEKV-GVELGKDVELNCSASLNKDDLFYWSIRKEDSSDPNVQEDRKETTTWIS 270

Qy         52 DGLPLGIGGHYSLHEYSWVKANLSEVLVSSVLGVNVTSTEVYGAFTCSIQN---ISFSSF 108
              :|                 | : |::|        :|   :   : |:: |   |   ||
Db        271 EG----------------KLHASKILRFQ----KITENYLNVLYNCTVANEEAIDTKSF 309

Qy        109 TLQR---AGPTSHVAAVLASLLVL--LALLLAALLYVKCRLNVLLWYQD-AYGEVEINDG 162
               |||        ||     ::||| :| :   :| |  :::::|:|:  |  :  : ||
Db        310 VLVRKEIPDIPGHVFTGGVTVLVLASVAAVCIVILCVIYKVDLVLFYRRIAERDETLTDG 369
```

```
Qy        163 KLYDAYVSY-SDC-PEDRKFVNFILK--PQ-LERRRGYKLFLDDRDLLPRAEPSADLLVN 217
              | |||:|||  :| ||:::   | ::  |: ||:: |||| : :||::|      ::
Db        370 KTYDAFVSYLKECHPENKEEYTFAVETLPRVLEKQFGYKLCIFERDVVPGGAVVEEIHSL 429

Qy        218 LSRCRRLIVVLSDAFLSRAWCSHSFREGLCR-LLELTRRPIFITFEGQRRDPAH------ 270
              : : ||||:||| ::|:           ||   |:|   : | | |        ||
Db        430 IEKSRRLIIVLSQSYLTNG-ARRELESGLHEALVERKIKIILIEF-----TPASNITFLP 483

Qy        271 PALRLLRQHRHLVTLLLWRPGSVTPSSDFWKEVQLALPRK 310
              |:|:||: :|    :| ||  | : :| ||| :   :| |
Db        484 PSLKLLKSYR----VLKWRADSPSMNSRFWKNLVYLMPAK 519


RESULT 13
ILAP_MACMU
ID   ILAP_MACMU     STANDARD;      PRT;   570 AA.
AC   P59822;
DT   10-OCT-2003 (Rel. 42, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Interleukin-1 receptor accessory protein precursor (IL-1 receptor
DE   accessory protein) (IL-1RAcP).
GN   Name=IL1RAP;
OS   Macaca mulatta (Rhesus macaque).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Macaca.
OX   NCBI_TaxID=9544;
RN   [1]
RP   SEQUENCE FROM N.A., FUNCTION, ALTERNATIVE SPLICING, AND
RP   CHARACTERIZATION OF ISOFORM 2.
RX   MEDLINE=22419343; PubMed=12530978; DOI=10.1016/S1074-7613(02)00514-9;
RA   Smith D.E., Hanna R., Friend D., Moore H., Chen H., Farese A.M.,
RA   MacVittie T.J., Virca G.D., Sims J.E.;
RT   "The soluble form of IL-1 receptor accessory protein enhances the
RT   ability of soluble type II IL-1 receptor to inhibit IL-1 action.";
RL   Immunity 18:87-96(2003).
CC   -!- FUNCTION: Mediates interleukin-1-dependent activation of NF-kappa-
CC       B. Isoform 1 is part of the membrane-bound form of the IL-1
CC       receptor. Signaling involves formation of a ternary complex
CC       containing IL1R1, TOLLIP, MYD88, and IRAK1 or IRAK2. Isoform 2
CC       modulates the response to interleukins by associating with soluble
CC       IL1R1 and enhancing interleukin-binding to the decoy receptor (By
CC       similarity).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein (isoform 1). Isoform
CC       2 is secreted.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=1; Synonyms=Membrane-bound IL-1R AcP;
CC         IsoId=P59822-1; Sequence=Displayed;
CC       Name=2; Synonyms=Soluble IL-1R AcP;
CC         IsoId=P59822-2; Sequence=Not described;
CC   -!- SIMILARITY: Belongs to the interleukin-1 receptor family.
CC   -!- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC   -!- SIMILARITY: Contains 1 TIR domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AY182233; AAO24704.1; -.
DR   EMBL; AY182234; AAO24705.1; -.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR004075; IL1_receptor1.
DR   InterPro; IPR004074; IL1_receptorI/II.
DR   InterPro; IPR000157; TIR.
DR   Pfam; PF00047; ig; 2.
DR   Pfam; PF01582; TIR; 1.
DR   PRINTS; PR01536; INTRLKN1R12F.
DR   PRINTS; PR01537; INTRLKN1R1F.
DR   SMART; SM00409; IG; 2.
DR   SMART; SM00255; TIR; 1.
DR   PROSITE; PS50835; IG_LIKE; 2.
DR   PROSITE; PS50104; TIR; 1.
KW   Alternative splicing; Glycoprotein; Immunoglobulin domain; Receptor;
KW   Repeat; Signal; Transmembrane.
FT   SIGNAL       1     20       Potential.
FT   CHAIN       21    570       Interleukin-1 receptor accessory protein.
FT   DOMAIN      21    367       Extracellular (Potential).
FT   TRANSMEM   368    388       Potential.
FT   DOMAIN     389    570       Cytoplasmic (Potential).
FT   DOMAIN      21    128       Ig-like C2-type 1.
FT   DOMAIN     136    226       Ig-like C2-type 2.
FT   DOMAIN     242    350       Ig-like C2-type 3.
FT   DOMAIN     403    549       TIR.
FT   DISULFID    47    114       Potential.
FT   DISULFID   137    181       Potential.
FT   DISULFID   160    212       Potential.
FT   DISULFID   266    332       Potential.
FT   CARBOHYD    57     57       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   107    107       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   111    111       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   118    118       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   196    196       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   209    209       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   299    299       N-linked (GlcNAc. . .) (Potential).
SQ   SEQUENCE   570 AA;  65393 MW;  9CD78AB1E78DD4C5 CRC64;

  Query Match             11.2%;  Score 240.5;  DB 1;  Length 570;
  Best Local Similarity   26.9%;  Pred. No. 1.6e-11;
  Matches   90;  Conservative   63;  Mismatches  144;  Indels   37;  Gaps   11;

Qy          9 PDFLSPSEDQVLRPALGSSVALNCTAWVVSGPHCSLPSVQWLKDG-----LPLGIGGHYS 63
              |   ||::  |      |  : : || :  |    |  | |  ||     :|: :  : |
Db        243 PVIHSPNDHVVYEKEPGEELLIPCTVY-FSFLMDSRNEVWWTIDGKKPDDIPIDVTINES 301

Qy         64 LHEYSWVKANLSEVLVSSVLGV-NVTSTEVYGAFTC---SIQNISFSSFTLQRAGPTSHV 119
              :       :   :   : :| :  ||| ::  :: |    | :     : |:::  |
Db        302 ISH-----SRTEDETRTQILSIKKVTSEDLKRSYVCHARSAKGEVAKAATVKQKVPAPRY 356

Qy        120 AAVLA---SLLVLLALLLAALLYVKCRLNVLLWYQDAYGEVE-INDGKLYDAYVSYSDCP 175
               ||      ||| ::|  : :|    | ::|:|:  :|  | | ||| || ||||:
Db        357 TVELACGFGATVLLVVILIVVYHVYW-LEMVLFYRAHFGTDETILDGKEYDIYVSYARNA 415

Qy        176 EDRKFVNFILKPQLERRRGYKLFLDDRDLLPRAEPSADLLVNLSRCRRLIVVLSDAFLSR 235
              |: :|   |:  ||   |||| : ||| ||    : : |  : : |||:|||  ::
Db        416 EEEEFVLLTLRGVLENEFGYKLCIFDRDSLPGGIVTDETLSFIQKSRRLLVVLSPNYV-- 473

Qy        236 AWCSHSFREGLCRLLELTRRPIFITFEG-------QRRDPAHPALRLLRQHRHLVTLLLW 288
                    :|   ||||      :  :|       | :      :: |:: : ::|:: |
Db        474 -------LQGTQALLELKAGLENMASQGNINVILVQYKAVKETKVKELKRAKTVLTVIKW 526

Qy        289 R-PGSVTPSSDFWKEVQLALPRKVRYRPVEGDPQ 321
              :   |  |   |||::|:|:| |   |    | |
Db        527 KGEKSKYPQGRFWKQLQVAMPVKKSPRRSSSDEQ 560


RESULT 14
IR18_HUMAN
ID   IR18_HUMAN     STANDARD;      PRT;   541 AA.
AC   Q13478;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Interleukin-18 receptor 1 precursor (IL1 receptor-related protein)
DE   (IL-1Rrp).
GN   Name=IL18R1; Synonyms=IL1RRP;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
```

```
RN   [1]
RP   SEQUENCE FROM N.A., AND VARIANT ALA-317 DEL.
RX   MEDLINE=96223957; PubMed=8626725; DOI=10.1074/jbc.271.8.3967;
RA   Parnet P., Garka K.E., Bonnert T.P., Dower S.K., Sims J.E.;
RT   "IL-1Rrp is a novel receptor-like molecule similar to the type I
RT   interleukin-1 receptor and its homologues T1/ST2 and IL-1R AcP.";
RL   J. Biol. Chem. 271:3967-3970(1996).
RN   [2]
RP   SEQUENCE OF 19-29; 55-58; 59-71; 211-221; 222-231; 269-275; 276-280;
RP   319-315 AND 524-535, AND CHARACTERIZATION.
RX   MEDLINE=97467370; PubMed=9325300; DOI=10.1074/jbc.272.41.25737;
RA   Torigoe K., Ushio S., Okura T., Kobayashi S., Taniai M., Kunikata T.,
RA   Murakami T., Sanou O., Kojima H., Fujii M., Ohta T., Ikeda M.,
RA   Ikegami H., Kurimoto M.;
RT   "Purification and characterization of the human interleukin-18
RT   receptor.";
RL   J. Biol. Chem. 272:25737-25742(1997).
CC   -!- FUNCTION: Receptor for interleukin 18 (IL-18). Binding to the
CC       agonist leads to the activation of NF-kappa-B.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- TISSUE SPECIFICITY: Expressed in lung, leukocytes, spleen, liver,
CC       thymus, prostate, small intestine, colon, placenta, and heart, and
CC       is absent from brain, skeletal muscle, pancreas, and kidney. High
CC       level of expression in Hodgkin's disease cell lines.
CC   -!- MISCELLANEOUS: Does not bind interleukin 1 alpha (IL-1A) or
CC       interleukin 1 beta (IL-1B).
CC   -!- SIMILARITY: Belongs to the interleukin-1 receptor family.
CC   -!- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC   -!- SIMILARITY: Contains 1 TIR domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U43672; AAC50390.1; -.
DR   HSSP; Q15399; 1FYV.
DR   Genew; HGNC:5988; IL18R1.
DR   MIM; 604494; -.
DR   GO; GO:0005886; C:plasma membrane; TAS.
DR   GO; GO:0004872; F:receptor activity; TAS.
DR   GO; GO:0006955; P:immune response; TAS.
DR   GO; GO:0007165; P:signal transduction; TAS.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR004075; IL1_receptor1.
DR   InterPro; IPR004074; IL1_receptorI/II.
DR   InterPro; IPR000157; TIR.
DR   Pfam; PF01582; TIR; 1.
DR   PRINTS; PR01536; INTRLKN1R12F.
DR   PRINTS; PR01537; INTRLKN1R1F.
DR   SMART; SM00409; IG; 2.
DR   SMART; SM00255; TIR; 1.
DR   PROSITE; PS50835; IG_LIKE; 1.
DR   PROSITE; PS50104; TIR; 1.
KW   Direct protein sequencing; Glycoprotein; Immunoglobulin domain;
KW   Polymorphism; Receptor; Repeat; Signal; Transmembrane.
FT   SIGNAL       1     18
FT   CHAIN       19    541       Interleukin-18 receptor 1.
FT   DOMAIN      22    329       Extracellular (Potential).
FT   TRANSMEM   330    350       Potential.
FT   DOMAIN     351    541       Cytoplasmic (Potential).
FT   DOMAIN      33    121       Ig-like C2-type 1.
FT   DOMAIN     133    212       Ig-like C2-type 2.
FT   DOMAIN     220    312       Ig-like C2-type 3.
FT   DOMAIN     373    523       TIR.
FT   DISULFID   140    185       By similarity.
FT   DISULFID   237    298       By similarity.
FT   CARBOHYD    91     91       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   102    102       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   150    150       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   197    197       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   203    203       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   236    236       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   255    255       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD   297    297       N-linked (GlcNAc. . .) (Potential).
FT   VARIANT    317    317       Missing.
FT                               /FTId=VAR_014955.
SQ   SEQUENCE   541 AA;  62304 MW;  7173DB9C7EA71D32 CRC64;

  Query Match             10.8%;  Score 231.5;  DB 1;  Length 541;
  Best Local Similarity   28.0%;  Pred. No. 8.5e-11;
  Matches   99;  Conservative   62;  Mismatches  126;  Indels   67;  Gaps   19;

Qy          6 DRA---PDFLSPSEDQVLRPALGSSVALNCTA---------WVVSGPHCSLPSVQWLKDG 53
              ||:   |  | |  : |     || :| |||:|          |:    : | |::   |:
Db        209 DRSNIVPVLLGPKLNHV-AVELGKNVRLNCSALLNEEDVIYWMFGEENGSDPNIHEEKEM 267

Qy         54 LPLGIGGHYSLHEYSWVKANLSEVLVSSVLG---VNVTSTEVYGAFTCSIQNISFSSFTL 110
                :   |          | : |:||    :|   :||    :|     |        || |
Db        268 RIMTPEG----------KWHASKVLRIENIGESNLNV----LYNCTVASTGGTDTKSFIL 313

Qy        111 QR----AGPTSHV--AAVLASLLVLLALLLAALLYVKCRLNVLLWYQDAYGEVE-INDGK 163
              |    |    ||     :: ::|:|:|::    : |  |::::|:|:      | : |||
Db        314 VRKADMADIPGHVFTRGMIIAVLILVAVVCLVTVCVIYRVDLVLFYRHLTRRDETLTDGK 373

Qy        164 LYDAYVSY-SDC----PEDRKFVNFILKPQLERRRGYKLFLDDRDLLPRAEPSADLLVNL 218
               |||:|||  :|      |:  |   ||    ||:  ||||  : :||::|       ::   :
Db        374 TYDAFVSYLKECRPENGEEHTFAVEILPRVLEKHFGYKLCIFERDVVPGGAVVDEIHSLI 433

Qy        219 SRCRRLIVVLSDAFLS---RAWCSHSFREGLCR------LLELTRRPIF-ITFEGQRRDP 268
               :|||:||| :::|   |       | |         |:| |  |:   ||  |
Db        434 EKSRRLIIVLSKSYMSNEVRYELESGLHEALVERKIKIILIEFT--PVTDFTFLPQ---- 487

Qy        269 AHPALRLLRQHRHLVTLLLWRPG-SVTPSSDFWKEVQLALPRKVRYRPVEGDPQ 321
               :|:||: ||    :| |:   |:: :| ||| :   :| |   :|   :|:
Db        488 ---SLKLLKSHR----VLKWKADKSLSYNSRFWKNLLYLMPAKT-VKPGRDEPE 533


RESULT 15
ILAP_HUMAN
ID   ILAP_HUMAN     STANDARD;      PRT;   570 AA.
AC   Q9NPH3; O14915; Q86WJ7;
DT   10-OCT-2003 (Rel. 42, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Interleukin-1 receptor accessory protein precursor (IL-1 receptor
DE   accessory protein) (IL-1RAcP).
GN   Name=IL1RAP;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 1), FUNCTION, AND INTERACTION WITH IL1R1
RP   AND IRAK1.
RC   TISSUE=Placenta;
RX   MEDLINE=98058729; PubMed=9371760; DOI=10.1073/pnas.94.24.12829;
RA   Huang J., Gao X., Li S., Cao Z.;
RT   "Recruitment of IRAK to the interleukin 1 receptor complex requires
RT   interleukin 1 receptor accessory protein.";
RL   Proc. Natl. Acad. Sci. U.S.A. 94:12829-12832(1997).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   TISSUE=Brain;
RA   Saito T., Seki N.;
RL   Submitted (AUG-1997) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), FUNCTION, AND INDUCTION BY
RP   PHORBOL ESTERS.
RC   TISSUE=Liver;
RX   MEDLINE=20261666; PubMed=10799889;
RA   Jensen L.E., Muzio M., Mantovani A., Whitehead A.S.;
```

```
RT   "IL-1 signaling cascade in liver cells and the involvement of a
RT   soluble form of the IL-1 receptor accessory protein.";
RL   J. Immunol. 164:5277-5286(2000).
RN   [4]
RP   SEQUENCE FROM N.A. (ISOFORMS 1; 2; 3 AND 4), AND INDUCTION BY PHORBOL
RP   ESTERS.
RC   TISSUE=Liver;
RX   MEDLINE=22664804; PubMed=12781872;
RA   Jensen L.E., Whitehead A.S.;
RT   "Expression of alternatively spliced interleukin-1 receptor accessory
RT   protein mRNAs is differentially regulated during inflammation and
RT   apoptosis.";
RL   Cell. Signal. 15:793-802(2003).
RN   [5]
RP   SEQUENCE FROM N.A. (ISOFORM 2).
RC   TISSUE=Skin;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [6]
RP   SEQUENCE OF 118-179 FROM N.A. (ISOFORM 1).
RX   MEDLINE=98140136; PubMed=9479509; DOI=10.1006/geno.1997.5113;
RA   Dale M., Hammond D.W., Cox A., Nicklin M.J.H.;
RT   "The human gene encoding the interleukin-1 receptor accessory protein
RT   (IL1RAP) maps to chromosome 3q28 by fluorescence in situ hybridization
RT   and radiation hybrid mapping.";
RL   Genomics 47:325-326(1998).
RN   [7]
RP   INTERACTION WITH IRAK2.
RX   MEDLINE=22546438; PubMed=12659850; DOI=10.1016/S0006-291X(03)00385-1;
RA   Boch J.A., Yoshida Y., Koyama Y., Wara-Aswapati N., Peng H., Unlu S.,
RA   Auron P.E.;
RT   "Characterization of a cascade of protein interactions initiated at
RT   the IL-1 receptor.";
RL   Biochem. Biophys. Res. Commun. 303:525-531(2003).
RN   [8]
RP   TISSUE SPECIFICITY.
RX   MEDLINE=22419343; PubMed=12530978; DOI=10.1016/S1074-7613(02)00514-9;
RA   Smith D.E., Hanna R., Friend D., Moore H., Chen H., Farese A.M.,
RA   MacVittie T.J., Virca G.D., Sims J.E.;
RT   "The soluble form of IL-1 receptor accessory protein enhances the
RT   ability of soluble type II IL-1 receptor to inhibit IL-1 action.";
RL   Immunity 18:87-96(2003).
CC   -!- FUNCTION: Mediates interleukin-1-dependent activation of NF-kappa-
CC       B. Isoform 1 is part of the membrane-bound form of the IL-1
CC       receptor. Signaling involves formation of a ternary complex
CC       containing IL1R1, TOLLIP, MYD88, and IRAK1 or IRAK2. Isoform 2
CC       modulates the response to interleukins by associating with soluble
CC       IL1R1 and enhancing interleukin-binding to the decoy receptor.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein (isoform 1). Isoform
CC       2, isoform 3 and isoform 4 are secreted.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=4;
CC       Name=1; Synonyms=Membrane-bound IL-1RAcP, mIL-1RAcP;
CC         IsoId=Q9NPH3-1; Sequence=Displayed;
CC       Name=2; Synonyms=Soluble IL-1RAcP, sIL-1RAcP;
CC         IsoId=Q9NPH3-2; Sequence=VSP_008050, VSP_008051;
CC       Name=3; Synonyms=Soluble IL-1RAcP-beta, sIL-1RAcP-beta;
CC         IsoId=Q9NPH3-3; Sequence=VSP_008052;
CC       Name=4;
CC         IsoId=Q9NPH3-4; Sequence=VSP_008053;
CC   -!- TISSUE SPECIFICITY: Detected in liver, skin, placenta, thymus and
CC       lung.
CC   -!- INDUCTION: Phorbol ester treatment causes down-regulation of
CC       isoform 1 and induction of isoform 2.
CC   -!- SIMILARITY: Belongs to the interleukin-1 receptor family.
CC   -!- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC   -!- SIMILARITY: Contains 1 TIR domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF029213; AAB84059.1; -.
DR   EMBL; AB006537; BAA25421.1; -.
DR   EMBL; AF167343; AAF71687.1; -.
DR   EMBL; AF167340; AAF71688.1; -.
DR   EMBL; AF167335; AAF71688.1; JOINED.
DR   EMBL; AF167336; AAF71688.1; JOINED.
DR   EMBL; AF167337; AAF71688.1; JOINED.
DR   EMBL; AF167338; AAF71688.1; JOINED.
DR   EMBL; AF167339; AAF71688.1; JOINED.
DR   EMBL; AF167342; AAF71689.1; -.
DR   EMBL; AF167335; AAF71689.1; JOINED.
DR   EMBL; AF167336; AAF71689.1; JOINED.
DR   EMBL; AF167337; AAF71689.1; JOINED.
DR   EMBL; AF167338; AAF71689.1; JOINED.
DR   EMBL; AF167339; AAF71689.1; JOINED.
DR   EMBL; AF167340; AAF71689.1; JOINED.
DR   EMBL; AF167341; AAF71689.1; JOINED.
DR   EMBL; AF538730; AAQ01755.1; -.
DR   EMBL; AF538731; AAQ01756.1; -.
DR   EMBL; AF538732; AAQ01757.1; -.
DR   EMBL; AF538733; AAQ01758.1; -.
DR   EMBL; AF538734; AAQ01759.1; -.
DR   EMBL; AF487335; AAO49451.1; -.
DR   EMBL; BC053621; AAH53621.1; -.
DR   EMBL; AF016261; AAC39609.1; -.
DR   Genew; HGNC:5995; IL1RAP.
DR   MIM; 602626; -.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR007110; Ig-like.
DR   InterPro; IPR004075; IL1_receptor1.
DR   InterPro; IPR004074; IL1_receptorI/II.
DR   InterPro; IPR000157; TIR.
DR   Pfam; PF00047; ig; 2.
DR   Pfam; PF01582; TIR; 1.
DR   PRINTS; PR01536; INTRLKN1R12F.
DR   PRINTS; PR01537; INTRLKN1R1F.
DR   SMART; SM00409; IG; 3.
DR   SMART; SM00255; TIR; 1.
DR   PROSITE; PS50835; IG_LIKE; 2.
DR   PROSITE; PS50104; TIR; 1.
KW   Alternative splicing; Glycoprotein; Immunoglobulin domain; Receptor;
KW   Repeat; Signal; Transmembrane.
FT   SIGNAL       1     20       Potential.
FT   CHAIN       21    570       Interleukin-1 receptor accessory protein.
FT   DOMAIN      21    367       Extracellular (Potential).
FT   TRANSMEM   368    388       Potential.
FT   DOMAIN     389    570       Cytoplasmic (Potential).
FT   DOMAIN      21    128       Ig-like C2-type 1.
FT   DOMAIN     141    230       Ig-like C2-type 2.
FT   DOMAIN     242    348       Ig-like C2-type 3.
FT   DOMAIN     403    549       TIR.
```

```
FT   DISULFID     47    114       Potential.
FT   DISULFID    137    181       Potential.
FT   DISULFID    160    212       Potential.
FT   DISULFID    266    332       Potential.
FT   CARBOHYD     57     57       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    107    107       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    111    111       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    118    118       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    196    196       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    209    209       N-linked (GlcNAc. . .) (Potential).
FT   CARBOHYD    299    299       N-linked (GlcNAc. . .) (Potential).
FT   VARSPLIC    351    356       VPAPRY -> GNRCGQ (in isoform 2).
FT                                /FTId=VSP_008050.
FT   VARSPLIC    357    570       Missing (In isoform 2).
FT                                /FTId=VSP_008051.
FT   VARSPLIC    302    570       ISHSRTEDETRTQILSIKKVTSEDLKRSYVCHARSAKGEVA
FT                                KAAKVKQKVPAPRYTVELACGFGATVLLVVILIVVYHVYWL
FT                                EMVLFYRAHFGTDETILDGKEYDIYVSYARNAEEEEFVLLT
FT                                LRGVLENEFGYKLCIFDRDSLPGGIVTDETLSFIQKSRRLL
FT                                VVLSPNYVLQGTQALLELKAGLENMASRGNINVILVQYKAV
FT                                KETKVKELKRAKTVLTVIKWKGEKSKYPQGRFWKQLQVAMP
FT                                VKKSPRRSSSDEQGLSYSSLKNV -> ASSKIHSGTGLWFW

  Query Match            10.6%;  Score 228.5;  DB 1;  Length 570;
  Best Local Similarity  27.6%;  Pred. No. 1.6e-10;
  Matches   91;  Conservative   60;  Mismatches  150;  Indels   29;  Gaps  11;

Qy          9 PDFLSPSEDQVLRPALGSSVALNCTAWVVSGPHCSLPSVQWLKDG-LPLGIGGHYSLHEY 67
              |    ||::  |      |  : : || :  |     |   | |  ||  |  |    :::|
Db        243 PVIHSPNDHVVYEKEPGEELLIPCTVY-FSFLMDSRNEVWWTIDGKKPDDITIDVTINE- 300

Qy         68 SWVKANLSEVLVSSVLGV-NVTSTEVYGAFTC---SIQNISFSSFTLQRAGPTSHVAAVL 123
              |   :   :   : :| :  ||| ::  :: |   | :     :  :::  |       |
Db        301 SISHSRTEDETRTQILSIKKVTSEDLKRSYVCHARSAKGEVAKAAKVKQKVPAPRYTVEL 360

Qy        124 A---SLLVLLALLLAALLYVKCRLNVLLWYQDAYGEVE-INDGKLYDAYVSYSDCPEDRK 179
              |       ||| ::|  : :|   | ::|:|:  :|  | | ||| || ||||:   |: :
Db        361 ACGFGATVLLVVILIVVYHVYW-LEMVLFYRAHFGTDETILDGKEYDIYVSYARNAEEEE 419

Qy        180 FVNFILKPQLERRRGYKLFLDDRDLLPRAEPSADLLVNLSRCRRLIVVLSDAFLSRAWCS 239
              ||   |:  ||   |||| : ||| ||    : : |  : : |||:||||  ::
Db        420 FVLLTLRGVLENEFGYKLCIFDRDSLPGGIVTDETLSFIQKSRRLLVVLSPNYV------ 473

Qy        240 HSFREGLCRLLELTRRPIFITFEG-------QRRDPAHPALRLLRQHRHLVTLLLWR-PG 291
              :|   ||||       :   |       | :      :: |:: : ::|:: |:
Db        474 ---LQGTQALLELKAGLENMASRGNINVILVQYKAVKETKVKELKRAKTVLTVIKWKGEK 530

Qy        292 SVTPSSDFWKEVQLALPRKVRYRPVEGDPQ 321
              |  |   |||::|:|:| |   |    | |
Db        531 SKYPQGRFWKQLQVAMPVKKSPRRSSSDEQ 560


Search completed: August  1, 2005, 13:49:22
Job time : 174 secs
```